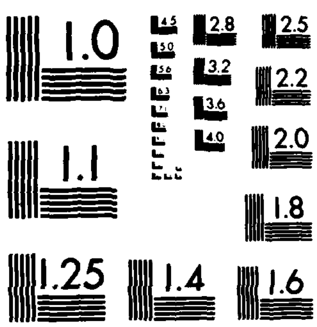
AD-A151 919 SURVEY OF NARROW BAND VOCODER TECHNOLOGY(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH W B MCMINN DEC 84 AFIT/CI/NR-85-24T 1/2

UNCLASSIFIED F/G 17/2 NL




MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

UNCLASS

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>AFIT/CI/NR 85-24T | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle)<br>Survey Of Narrowband Vocoder Technology | | 5. TYPE OF REPORT & PERIOD COVERED<br>THESIS/~~DISSERTATION~~ |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>William Boyce McMinn, Jr. | | 8. CONTRACT OR GRANT NUMBER(s) |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>AFIT STUDENT AT: Mississippi State Univ | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>AFIT/NR<br>WPAFB OH 45433 | | 12. REPORT DATE<br>Dec 1984 |
| | | 13. NUMBER OF PAGES<br>163 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | | 15. SECURITY CLASS. (of this report)<br>UNCLASS |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

APPROVED FOR PUBLIC RELEASE: IAW AFR 190-1

Lynn Wolaver 28 Feb 85
LYNN E. WOLAVER
Dean for Research and
Professional Development
AFIT, Wright-Patterson AFB OH

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

DTIC
ELECTE
MAR 27 1985
S D
E

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

ATTACHED

DD FORM 1473 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASS

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

85 03 11 044

24

# SURVEY OF NARROWBAND VOCODER TECHNOLOGY



By

WILLIAM BOYCE McMINN, JR.



A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in the Department of Electrical Engineering

Mississippi State, Mississippi

December 1984

ABSTRACT

The USAF is trying to identify a vocoder to insert into a Low Probability of Intercept (LPI) communications system. It should be small, light weight, low power, and capable of processing intelligible, natural sounding speech at 400 to 600 b/s. Two separate systems are needed: one to be utilized soon in a brassboard system to test the LPI concepts and one to be available as mid 90s off-the-shelf hardware for a production LPI system.

Weighted characteristic values are combined through a mapping and summing procedure to form a Figure of Merit, $F_s$, for comparing the systems. Each characteristic has a minimum, below which the system is considered unacceptable. Thirty eight current systems or research efforts were identified. Of these, only 17 were determined to be available in the desired time frame. These separated into 2 groups: mid 80s available and mid 90's available systems. The systems in each category were compared. The 3 with the highest $F_s$ values were identified as the primary candidates. For the mid 80s effort the optimum systems are all Motorola prototypes. They are: (1) the Miniaturized Narrowband Secure Voice System, (2) the Manpack Vocoder, and (3) the Advanced Technology Model Multi-Rate Processor LPC Vocoder. For the mid 90s effort none of the systems met all of the minimum requirements. The desired data rates and equipment sizes have not been combined in a single effort. More R&D funds are necessary to advance the development of vocoders to the desired stage.

SURVEY OF NARROWBAND VOCODER TECHNOLOGY

By

WILLIAM BOYCE McMINN, JR.

APPROVED:

Frank Ingels

Professor of Electrical Engineering
Major Professor

Frank Ingels

Professor of Electrical Engineering
Graduate Coordinator

Walter R. Carnes

Director of Graduate Instruction
College of Engineering

Willie L. McDaniel Jr.

Dean of the College of Engineering

Marvin T. Loftin

Vice-President for Graduate Studies
and Research

## PREFACE

The work reported in this thesis was performed at the United States Air Force Avionics Laboratory, a center for research operated by the Air Force Wright Aeronautical Laboratories (AFWAL) with the support of the Department of the Air Force.

The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the United States Air Force or the United States Government.

The value judgements of the specific equipments presented in this paper are a result of the author's choice and structuring of the evaluation criteria. The results of the comparisons are not intended to and do not reflect the quality or operational competance of any product discussed.

ACKNOWLEDGEMENT

I wish to express my sincere thanks to Mr. Lawrence L. Gutman who provided this topic for my thesis research and who has assisted me for about six months in the preparation of the final document. Additionally, I wish to thank the personnel of AFWAL/AAAI who put up with me for those six months, their assistance is greatly appreciated.

During the course of this research, I talked with several research engineers and company marketing personnel. These people are too numerous to thank here individually but they have my deepest appreication. I would especially like to thank the personnel of RADC/EEV, Hanscom AFB, Massachusetts, and MIT's Lincoln Laboratories also at Hanscom AFB for providing a tremendous amount of information and the opportunity to visit their facilities to discuss vocoders and to view ongoing experimentation in vocoder development.

Special appreciation goes to Professor Frank M. Ingles of the Electrical Engineering Department at Mississippi State University. Professor Ingles provided a large amount of initial guidance and suggestions and then let me pursue this project on my own.

Finally, I wish to thank my wife and family for putting up with the extended separations and numerous hours spent traveling back and forth between Wright-Patterson Air Force Base, Ohio, and Mississippi State University, Mississippi State, Mississippi.

# ABSTRACT

William B. McMinn, Jr., Master of Science, December 1984.

Major: Electrical Engineering, Department of Electrical Engineering.

Title of Thesis: Survey of Narrowband Vocoder Technology

Directed By: Dr. Franklin M. Ingels

Pages in Thesis: 163 Words in Abstract: 113

## Abstract

The USAF has a need to identify a vocoder to insert into a Low Probability of Intercept (LPI) communications system. It should be small, lightweight, low power, capable of operating in many types of aircraft, and capable of processing intelligible, natural sounding speech at 400 to 600 bits/seconds. Two separate units are needed: one to be used in a near-term brassboard test system and one to be used in a far-term production system. Weighted characteristic values are combined through a mapping and summing procedure to form a Figure of Merit for each system. Using these characteristic values, primary vocoder candidates have been identified and are discussed in this paper. Additional keywords: waveforms, theses, state of the art, brassboard models, multimode, [illegible] (LPI Antijam Secure Airborne [illegible] System).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# CHAPTER I

# INTRODUCTION

## Problem Statement

The United States Air Force is investigating the feasibility of implementing a Low Probability of Interception (LPI) communications system. The objective of this report is to provide an independent, non-government determination of what the current level in vocoder technology is and to evaluate it for application to the USAF LPI Communications Techniques Advanced Development Program (LPI Comm ADP).

The objective of the LPI Comm ADP is to develop from "off-the-shelf" technology a flight qualifiable brassboard system for advanced development testing in the late 1980s. The brassboard development is aimed at a mid-1990s production of a multimode LPI/Anti-Jam/Secure Airborn Radio System (LASARS).

At the present time, a conceptual design study for an LPI communications system is being performed under contract for the Air Force. This study is examining the feasibility of utilizing advanced device technology and is looking to integrate potential LPI signal processing techniques into the LARSARS. The technologies which yield LPI gains and are under consideration include spread spectrum modulation techniques, continous adaptive power control, speech bandwidth compression, adaptive interference suppression, adaptive beam pointing

antennas, low side lobe antennas, adaptive frequency control (multiple-band operations), adaptive null-steering antennas, and adaptive signal masking. Figure 1-1 depicts the LPI system in general terms and indicates where the various technologies fit within the projected system. The J's and I's indicate jamming and interference signals, respectively. The ESM block is the electronic support measures of the jamming and/or intercepting receivers.

This report will address the speech bandwidth compression area of the ongoing investigation by evaluating vocoder technology applicable to both the 1980s brassboard and the 1990s multimode LASARS. The report will be utilized by the LPI Comm ADP program manager and by the conceptual design contractor in their analysis of and recommendation for an overall communication system design integrating the LPI technology areas.

## Scope

This report provides a comprehensive overview of current vocoder production systems and of current research models for future implementations. It contains discussions in several areas. These include:

1. A description of the speech wave and the problems associated with characterizing it.

2. An analysis of LPI gains from reduced information data rates.

3. A summary of the theory of operation of each competitive vocoder technology.

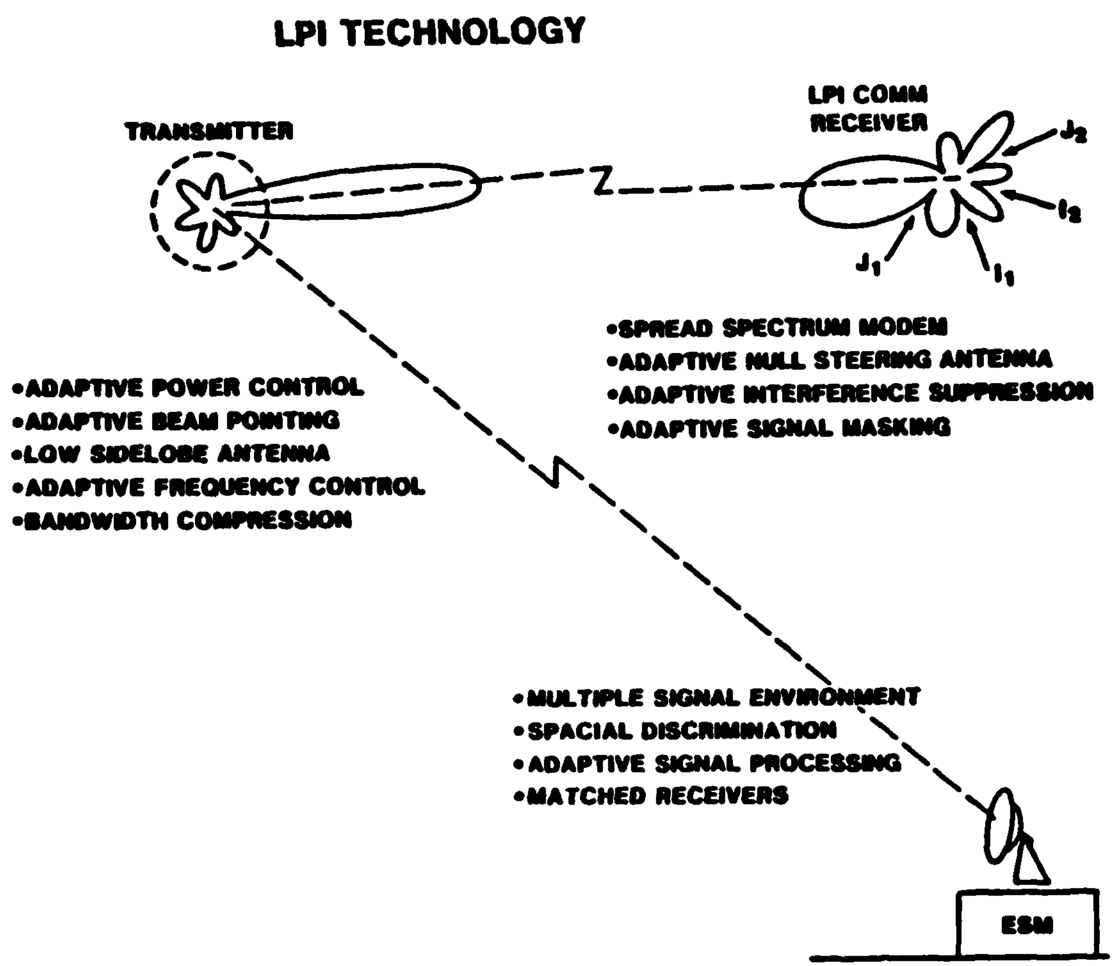



Figure 1-1. LPI technology conceptual design areas.
(Drawn from USAF/AFWAL VU-graphs.)

4. The development of an assessment criterion for comparing and selecting the most satisfactory vocoder configuration for near-term brassboard feasibility tests and for far-term LASARS implementation.

5. A qualitative comparison of low-rate (narrowband) vocoder technologies.

6. A quantitative comparison of low-rate vocoder technologies.

This report does not include any development of new vocoder technology or any new methods for analyzing vocoders. Its purpose is strictly to summarize the state of the art of vocoder technology, including all alternate approaches, and specific vocoder parameters and to select the vocoder candidates best suited for insertion into a brassboard and for application to a LASARS development and implementation.

## Approach

This report is generated from an extensive literature search. The data sources include a periodical search and a government and civilian technical report search conducted by accessing Department of Defense (DoD), Air Force, National Aeronautics and Space Administration (NASA), and various civilian report documentation data bases, by conducting library searches in three different libraries, by conducting private communications with engineers at various research organizations and private companies, and by visiting some research organizations.

The report develops the mathematical relationship between data bandwidth and the required transmitter power (Chapter II). This

chapter also develops an understanding of the problems associated with analyzing the speech waveform and the basic theory of operation of the major vocoder methods. A method of assessment for comparing and selecting a specific vocoder is developed (Chapter III). The report performs an in depth qualitative (Chapter IV) and an in depth quantitative (Chapter V) comparison of the various vocoder systems being produced and/or researched. Finally, it presents a set of recommendations (Chapter VI) concerning the best vocoder implementations for brassboard and production LPI communications systems.

## CHAPTER II

## BACKGROUND

### Overview

Before discussing vocoder systems and implementations it is necessary to understand the concept of an LPI communications system/channel and what gains are to be achieved in this channel through the use of vocoders. Also necessary is a discussion of the speech waveform characteristics and how they must be approached for an in depth mathematical analysis and characterization. Finally, the various vocoder methods or technologies must be presented to form a basis for examining specific implementations.

### Function of Vocoder

The vocoder will constitute a major function of the communications link. A typical digital communications system block diagram is shown in Figure 2-1(a). The vocoder will form the basis for the formatting/source encoding portion of the system. Figure 2-1(b) summarizes the typical functions of the various portions of the system. Different speech digitizers can be described as follows (1).

> Speech coders can be divided into two broad categories: waveform coders and vocoders. The waveform coders attempt to mimic the speech waveform as closely as possible. These coders are capable of producing high-quality speech but only at bit rates

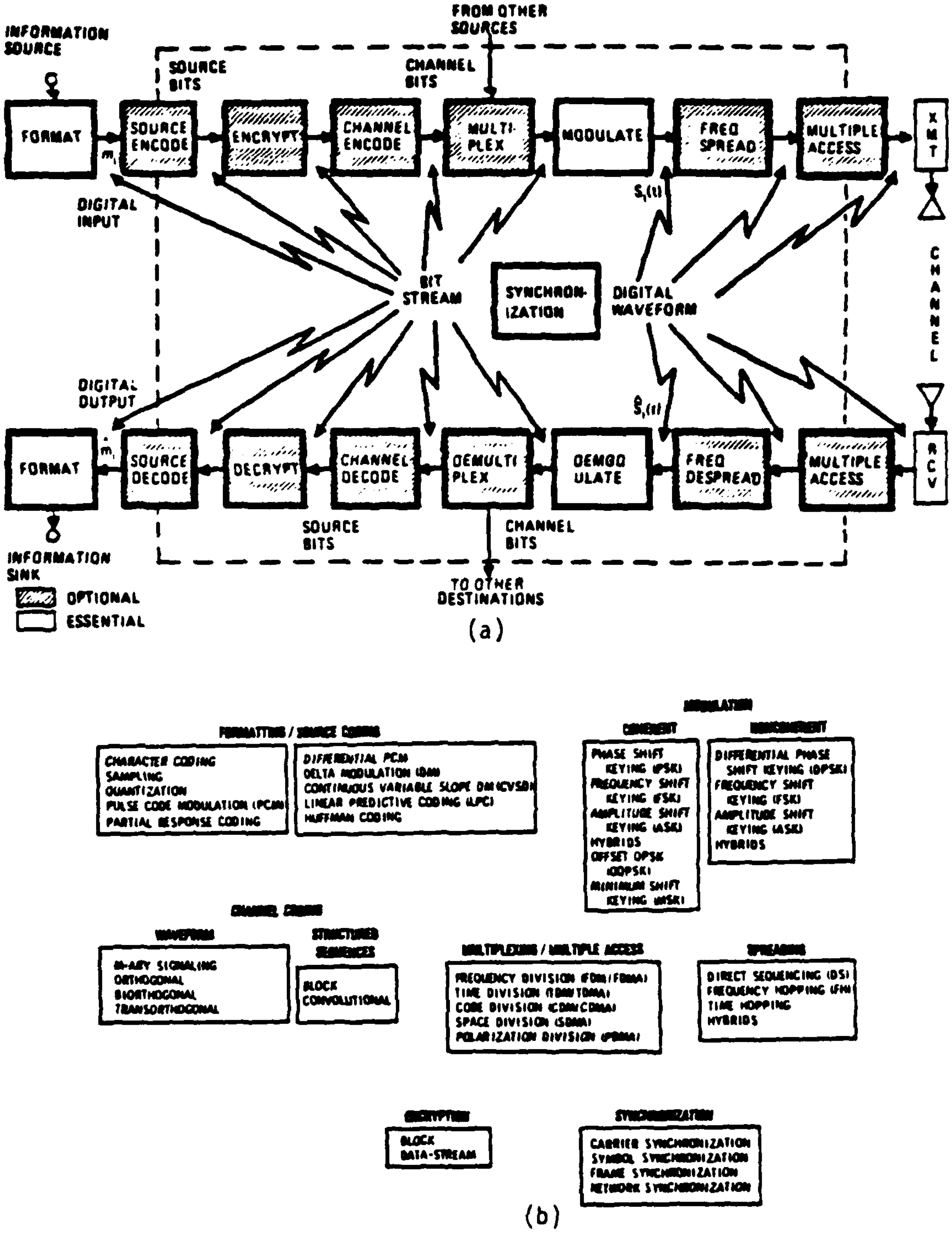



Figure 2-1. Digital communication system; (a) typical block diagram and (b) basic digital communication transformations. (Taken from ref. 54)

above about 16 kbits/s; the speech quality deteriorates significantly as the bit rate is lowered below 10 kbits/s. Vocoders are a parametric model of human speech production to achieve coding efficiency. Vocoders can produce intelligible speech but often the output speech has a synthetic quality. The speech quality from vocoders cannot generally be improved by increasing the bit rate.

The vocoder samples and quantizes the analog speech signal. It then performs one of a variety of mathematical analysis techniques in order to characterize the speech wave. This characterization is then output in digital form from the source encode portion of the system. The vocoder directly determines the basic data rate of the system. Encryption, when utilized, usually does not add to the data rate. When providing data error protection, the channel encoder adds bits proportional to the incoming data rate. The output power required to maintain a specific signal-to-noise ratio is directly proportional to the system information data rate. Therefore, decreasing the data rate will allow immediate reductions in output power allowing the system to operate at a minimum level and decreasing the detectability of the communications signal.

The system data rate is directly proportional to the system bandwidth which determines the necessary power output. Reducing the data rate allows corresponding reductions in the bandwidth. Reductions in bandwidth limit the input noise power. If the noise power input is reduced, the signal-to-noise ratio increases. Therefore, the transmitter output power can be reduced in order to maintain a given, required signal-to-noise ratio. Again, the system is assumed to be operating at a minimum level of acceptable communications. It is desired therefore to reduce the data rata as low as possible while

maintaining an acceptable level of speech recognition and quality and acceptable speaker identification. Discussion of the specific vocoder gains is presented in Appendix A1.

## Speech Waveform Considerations

The speech waveform presents unique problems for mathematical analysis. It has three general characteristics to be determined. It consists of "voiced" sounds, "unvoiced" sounds, and a basic pitch period. Voiced sounds are more-or-less periodic in nature and are almost the same in shape for every individual speaker with some slight differences in frequency content. Unvoiced sounds are noise-like in nature, varing only in frequency content and amplitude. The pitch period determines the basic repetition period for the voiced sounds and is a function of the vocal tract and different for every speaker. Figure 2-2 shows samples of these characteristics. Figure 2-3 shows samples of the frequency content of these two speech types. Speech production is generally characterized as the convolution of a given excitation, pitch pulses, with the vocal tract impulse response.

The speech production process is best described in the following quote (99).

> The human speech production system consists of an air pressure source (the lungs) feeding through the vocal cords and combined oral and nasal passages. The vocal cords can be caused to vibrate and provide a periodic (voiced) excitation to the vocal tract, or to be abducted to allow airflow into the tract. A constriction in the tract during a period of airflow can cause turbulence (friction) noise to be generated just downstream of the constriction (with or without voicing). The oral and nasal passages form a variable configuration deformable set of resonators connected to the excitation sources. Radiation of the resulting pressure wave from the mouth and nose causes the speech. Although the sources are nonwhite and the radiation process is frequency dependent, the model can be simplified by combining the frequency dependent

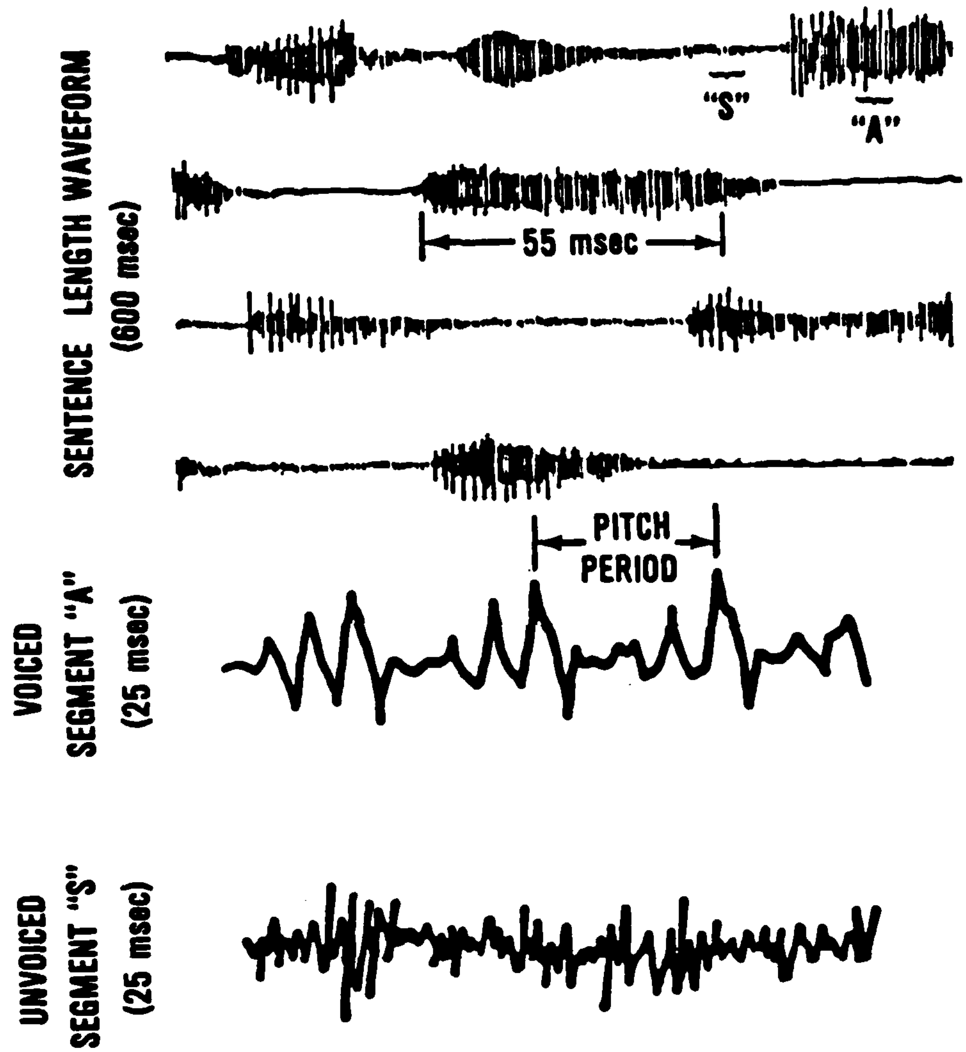



Figure 2-2. Examples of "long time" (sentence length) and "short time" (sound length) speech segments.
(Taken from reference 19, page 3G.3.1)

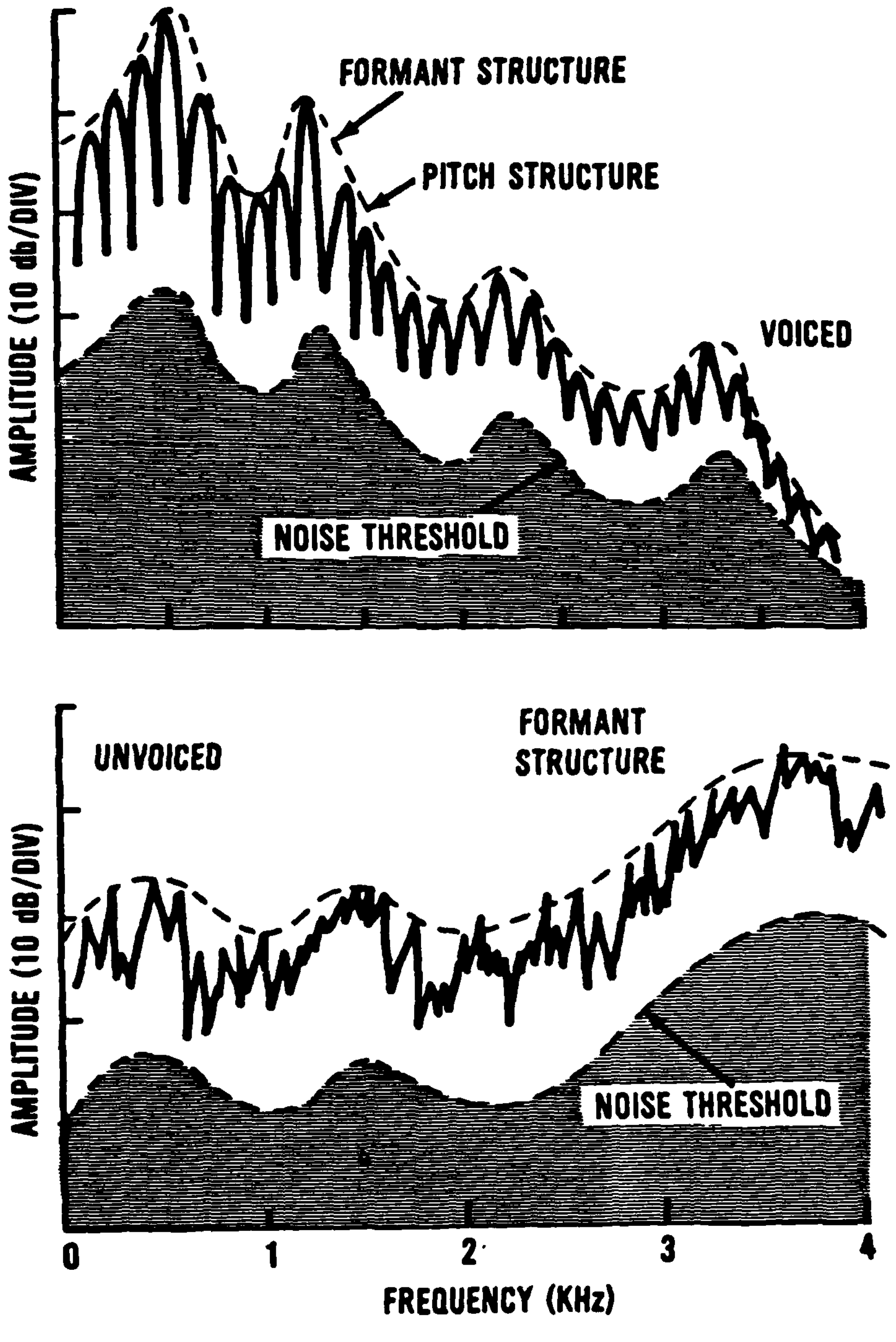



Figure 2-3. Spectral models for voiced and unvoiced speech. (Taken from reference 19, page 3G.3.1)

effects into a single filter that contains both poles and zeros, and by assuming that the source is either a periodic impulse train or white noise.

During the formation of a word or syllable several of these voiced and unvoiced sounds are usually combined. This means that the speech wave is constantly changing. In order to analyze the speech wave it is normal to "window" the speech signal. A specific time segment of the wave is chosen, usually around 20 milliseconds, during which the speech signal is considered to be stationary. Different analysis methods have different optimum window times determined through extensive research. A discussion of this determination is beyond the scope of this paper. A more detailed presentation of the characteristics of the speechwave can be found in Appendix A2.

## Vocoder Methods

Speech signals are known to be highly redundant. The most effective method to reduce this redundancy and thus reduce the channel capacity required for transmission of speech signals is to extract and transmit only the major characteristics of speech at a regular interval (typically 10 to 30 ms). These characteristics are then used at the receiver to reconstruct speech. A system based on this method is commonly known as an analysis-synthesis system. (132)

Analysis-synthesis systems are also known as vocoders. These are systems for analyzing, parameterizing, quantitizing, and then resynthesizing the speech waveform. These vocoders constitute a class of speech bandwidth compression techniques described in general terms by the following quote. (99)

Narrow-band speech compression systems generally require analysis and synthesis of speech with separate characterizations of the vocal tract frequency response and the excitation. These systems, generically termed vocoders, achieve reasonable quality speech reproduction at rates of 5 kbits/s and below by transmitting quantized parameters of this source-filter model. This bandwidth

reduction is generally achieved at the cost of a loss of voice intelligibility and naturalness. In addition, vocoder performance tends to be talker-sensitive, and to be fragile in that extraction of source and vocal tract parameters are affected adversely by additive acoustic noise and signal distortion at the vocoder input.

There are seven major, distinct voice analysis-synthesis methods. These are the channel, formant, homomorphic, pattern-matching, phase, linear predictive coding, and spectral envelope estimation vocoders. Other methods exist, none in production or undergoing extensive research, which combine aspects of the various methods. These combination methods will not be considered.

Channel Vocoder. The channel vocoder or spectrum-channel vocoder separates the speech signal into 12 to 28 frequency channels. Contiguous bandpass filters are used to perform this separation. Each channel is rectified and then low pass filtered. The time-varying signal then represents the amount of signal energy in the given frequency range. This can then be quantized with a few bits and transmitted. A final channel consists of a voiced/unvoiced (V/UV) detector and a pitch extractor. This information is quantized and transmitted and then used in the synthesizer along with the channel signals to control the frequency response of a time-varying resonant filter to correspond to the spectral envelope measured at the analyzer. See Appendix A3.

Formant Vocoder. The formant method assumes that the speech wave can be characterized by an envelope consisting of several prominent maxima. Below 3,000 Hz there are usually three maxima and below

4,000 Hz there are four or five maxima. These are known as "formants." Here the analyzer determines the frequency location, bandwidths, and relative amplitudes of the individual formants. This information is coded and transmitted to the synthesizer which uses it to control resonances of a formant synthesizer consisting of tuned resonant circuits. There are several methods of determining the formant frequencies. One method is to measure the rate of axis crossing of filter separated formants. This method is fairly inaccurate without additional adjustments. A second method is to "channelize" the signal, measure the amplitude of each channel, and average to determine the most prominent frequencies. A final major method involves finding the average of the derivative of the time signal and dividing by the average of the time signal itself. A prominent vocoder method is an analysis-by-synthesis vocoder. It is one form of a formant vocoder. It generates known artifical spectra and compares them to the incoming speech wave and through iteration methods these spectra are matched and the known characteristics of the "artificial" signal are transmitted. The pitch extraction and V/UV decision is made the same as in the channel vocoder. See Appendix A3.

Homomorphic Vocoder. The homorphic or cepstrum vocoder utilizes an FFT approach through the use of homomorphic filtering concepts. The convolved speech signal is transformed into a spectral magnitude, product signal by a high-resolution Fourier transform. This is transformed into an addition process by taking the logarithm of the spectral magnitude. This yields a rapidly varying pitch component and a slowing-varying vocal tract component. Now, another

Fourier transform (or inverse transform) [95,104] is performed separating the signal into a "low-time" component containing the vocal tract information and a "high-time" component containing the excitation or vocal-cord information. Pitch extraction and the V/UV decision is determined utilizing the high-time information. This method is popular for use with other vocoder methods as the pitch extraction and V/UV decision determination method. See Appendix A3.

Pattern Matching Vocoder. The pattern-matching vocoder transmits basically three items of information. It transmits the memory location of a stored spectral pattern which most closely matches the speech segment, the pitch information, and the V/UV decision. Sometimes error information about the difference between the stored pattern and the speech segment is sent so that it can be used to adjust the pattern recalled from memory. This increases the data rate somewhat and most designers have determined this information to be essentially unnecessary. Almost any speech analysis method can be utilized to transform the speech signal into a form matching those in memory. Windowed speech segments are again used. With today's high speed computers smaller windows are used allowing more comparisons and more accurately synthesized speech.

Phase Vocoder. The phase vocoder utilizes a channel method to generate "short-time" or windowed amplitude and phase spectra to represent the characteristics of the speech wave. This method differs from the channel vocoder in that the derivative of the signal phase is determined in the analysis procedure and utilized in the synthesis

procedure to regenerate the speech signal. With this method no pitch or V/UV information is needed to modulate with the spectral information in the regeneration. See Appendix A3.

Linear Prediction Vocoder. Linear predictive coding vocoders operate in the time domain rather than in the frequency domain. This method uses weighted sums of a given number of past samples in order to predict the present sample. The weights form the adaptive portion of the analysis. They are adjusted to minimize the error signal between the actual and predicted speech samples. The system transmits selected characteristics of the error signal. These transmitted signals include predictor coefficients, gain, pitch information, and the V/UV decision. In the process of determining the predictor coefficients several intermediate sets of coefficients are formed. Any of this information can be transmitted with the most common set being the reflection coefficients. See Appendix A3.

Spectral Envelope Estimation Vocoder. A final major vocoder or analysis-synthesis method is the spectral envelope estimation vocoder. This method is the most recently developed technique. In addition to generating parameters used in the regeneration of the speech signal this method provides an estimate of the background noise for use in noise suppression. The system approximates the spectral envelope (the vocal tract filter response) of the speech wave. The pitch extraction technique forms an intergral part of the analysis system. The average pitch is constantly fed back into the system so that peaks in the speech wave can be estimated. The smoothing of these peaks form the estimated envelope to be quantized for

transmission. The pitch and V/UV information is also transmitted. This method, as in the previous methods, uses a windowed frame of speech but this window is adaptive in length, averaging about 2.5 times the pitch period in length. The background noise estimator, when used, adds to the vocoder data rate. See Appendix A3.

In all methods, except the phase vocoder, pitch and voicing information is necessary. There are very many different methods for extracting this information [4, 16, 34, 42, 56, 78, 95, 99, 115]. Generally, the periodic nature of voiced sounds is detected from the voiced signal. If no periodic or very slowly varying set of peaks is detected, the unvoiced signal is generated. Only one bit is needed for transmission to indicate the V/UV decision. The detected periodic signal is measured to determine the pitch period. Usually about seven bits are used to quantize this parameter.

The pitch extraction--V/UV decision process is an integral part of any vocoder. Therefore, when selecting a particular vocoder no choice of this process is available. The designer of the vocoder made the choice as to which pitch extraction algorithm best works with or is most economical in the vocoder designed. Further information on pitch extraction--V/UV decisions can be found in almost every reference listed in the Bibliography.

# CHAPTER III

# ASSESSMENT METHODOLOGY

## Overview

Vocoder systems are implemented in a variety of methods as previously mentioned in Chapter II. Additionally, each method usually has a number of different implementations. This diversity of systems establishes the need for a method to quantitatively compare these systems with each other. This chapter presents parameters, the parameter constraints, and the minimum specifications used for the vocoder evaluation. A Figure of Merit analysis method is developed for comparing the vocoders. Then the differences in evaluating a system for near-term and far-term implementations are discussed with the associated modifications in the Figure of Merit analysis presented. Finally, a discussion of each parameter and the choice of the minimum specifications is presented.

## Parameters and Constraints

Any piece of equipment can be compared in terms of performance, size, weight, cost, availability, etc. When parameters of this type are considered, minimum acceptable values are usually assigned. In determining these minimum values, several factors interact to provide constraints. The constraining factors in the selection of a

vocoder to be used on board an aircraft are: (1) mission requirements, (2) platform requirements (aircraft physical limitations), and (3) user requirements. Quite often state-of-the-art technology capabilities require that some parameter constraints be relaxed. Minimum values assigned a parameter can be a result of one or more of the constraints mentioned above. Table 3-1 lists those parameters considered important in evaluating a vocoder for an airborne brassboard LPI radio system. The table lists the parameters in order of decreasing importance. Included in the table is (are) the constraining element(s) for each parameter (identified as 1, 2 or 3 from above) and the associated minimum specification.

### Figure of Merit Analysis Near-Term Brassboard System

The brassboard system conceptual design testing will be initiated within two or three years from the present. Based upon experience (54, 120, 124, 126, 136) with hardware development, this means the hardware must be available now either in production or as an engineering prototype which a company would be willing to sell to the Air Force. This fact alone eliminates about 80 percent of the vocoder systems/methods being researched by various organizations. For this phase of the LPI Comm ADP only one or two systems will need to be purchased for testing purposes. The flight tests will be conducted to determine the feasibility of implementing an LPI communciations system. Each of the parameters/specifications listed in Table 3-1 can take on a range of values given in the Figure of Merit Analysis chart shown in Table 3-2. These values are listed in a set of columns

TABLE 3-1

SYSTEM PARAMETERS WITH CONSTRAINTS AND
MINIMUM SPECIFICATIONS FOR
NEAR-TERM ANALYSIS

| Parameter | Constraints* | Specification |
|---|---|---|
| Data Rate (bit per second or b/s) | 1 | $\leq$ 2400 |
| Intelligibility | 1, 3 | $\geq$ 80% DRT† |
| Input Probability of Error | 1 | $\geq 10^{-3}$ |
| Size | 2 | $\leq$ 3000 $in^3$ (1.76 $ft^3$) |
| Weight | 2 | $\leq$ 50 lb. |
| Power Consumption | 2, 3 | $\leq$ 100 W |
| Processing Delay | 1, 3 | Real time,(< 100 ms throughput) |
| System Availability | 1, 3 | Engineering Prototype (minimum) |
| Production Cost (Lots of 1,000) | 3 | $\leq$ $40,000/unit |
| Speaker Dependence | 3 | None |
| Vocabulary Dependence | 3 | None |
| System "Learning" Time | 1, 3 | None |

* 1. Mission requirements
2. Platform requirements
3. User requirements

† Diagnostic Rhyme Test Score (discussed later)

TABLE 3-2

FIGURE OF MERIT ANALYSIS SYSTEM FOR BRASSBOARD INSERTION SYSTEM (Sample Chart)

SYSTEM: ______________________________

SOURCE: ______________________________

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | > 2400 | | |
| INTELLIGIBILITY (% DRT) | .225 | > 95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | | |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | | |
| SIZE ($in^3$) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | > 3000 | | |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | | |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | > 100 | | |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | | |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | | |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | > 40k | | |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | | |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | | |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | | |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | |

labeled "Parameter Figure of Merit." Each of the values listed are mapped with a one-to-one correspondence to the Figure of Merit ranging from 0 to 10 heading each column resulting in an assigned figure of merit, Fp. In the table, each parameter is mapped to a weight, Wp, according to the relative importance of the parameter. The sum of the weights is normalized to 1.0. The product of the parameter weight and the parameter Figure of Merit results in the parameter score, $P_s$,

$$P_s = W_p \times F_p \quad . \tag{3-1}$$

the sum of the parameter scores is the System Figure of Merit, $F_s$,

$$F_s = \sum_{i=1}^{12} P_{s_i} \quad , \tag{3-2}$$

ranging in value from 0.000 to 10.000. Each system under consideration is evaluated independently on a table identical to the one presented here. After system evaluation, the systems are compared to each other using the system Figure of Merit totals. The system having the highest Figure of Merit is, theoretically, the optimum system.

In the mapping of parameter values to Figures of Merit, it is not necessary that a linear relationship exist. For example, the data rate mapping of b/s to the Figure of Merit is shown in Table 3-3.

TABLE 3-3

DATA RATE MAPPING

| Parameter Value | <-----> | Figure of Merit |
|---|---|---|
| <150 | | 10 |
| 150 | | 9 |
| 200 | | 8 |
| 300 | | 7 |
| 400 | | 6 |
| 600 | | 5 |
| 800 | | 4 |
| 1,200 | | 3 |
| 1,600 | | 2 |
| 2,400 | | 1 |
| >2,400 | | 0 |

Examination of Table 3-2 shows that this type of nonlinearity exists for most parameters. Not all of the parameters have a range of 11 values. In these cases, the range is distributed as evenly as possible over the Figure of Merit mapping range. This is best demonstrated by the availability mapping for systems under consideration for brassboard insertion. This is detailed in Table 3-4.

In each system evaluation according to Table 3-2, any parameter scoring a 0 parameter Figure of Merit is marked with an asterisk (*) in the parameter score column. Any system marked in this manner has fallen below a minimum specification and is therefore deleted from further consideration. The data on these systems is still provided in the event that future considerations dictate a relaxing of any minimum specifications.

TABLE 3-4

DATA RATE MAPPING

| Parameter Value | <-----> | Figure of Merit |
|---|---|---|
| Production | | 10 |
| ----- | | 9 |
| ----- | | 8 |
| ----- | | 7 |
| ----- | | 6 |
| Engineering Prototype | | 5 |
| ----- | | 4 |
| ----- | | 3 |
| ----- | | 2 |
| ----- | | 1 |
| All Others | | 0 |

Quite often in evaluating the systems, one or more parameter values may not be available. In this case, two methods for adjusting the table exist. In the first method, the parameter value can be estimated by conversation with the developing engineers or by comparison with similar systems. If estimates are not possible, the parameter can remain unassigned and the system figure of merit is renormalized, $\bar{F}_s$, by dividing it by the sum of the parameter weights which are assigned,

$$\bar{F}_s = \frac{F_s}{\sum_i^N W_{p_i}} \quad . \tag{3-3}$$

In either case, the system involved is flagged so that the Air Force contractor or investigator can identify it and attempt to get more accurate data if desired. In this evaluation, both methods will be

utilized with estimated systems flagged with a capital E and renormalized systems flagged with a capital R.

Figure of Merit Analysis, Far-Term LASARS

In the event the brassboard flight tests indicate that LPI communications constitute a viable concept and given a continuing need with the associated funds appropriated, the Air Force will proceed into a system development phase. When the decision to continue is made, a final decision as to which component subsystems to purchase and incorporate will be made. This implementation is projected to occur in 1993 or 1994. At this time, all of the subsystems must be through engineering development and military specification testing and be ready for production with the producing company already setting up their production line. Based on the group experience mentioned previously, the engineering development requires three to five years with the following military specification testing requiring an additional two to five years. In the worst case, this means ten years of development and testing are required with additional time required to establish production capabilities. In order to meet this deadline, the systems to be considered must currently exist as laboratory models or as algorithm simulations designed to modify existing systems. Totally new techniques, now existing only as computer simulations, will probably require twelve to fifteen years to reach the stage required by the Air Force (54).

Current experiments with modifications in quantizing schemes and new insights into perceptual differences indicate that lower data

rates with improved performance will be available. Additionally, new systems implemented with the state-of-the-art technology in Very Large Scale Integration (VLSI) and in Very High Speed Integrated Circuits (VHSIC) will be smaller, lighter, and require less power than current production equipment. This leads to a tightening of the minimum specifications. Added to these reductions are a tightening of the requirements as a result of the applications of the system being designed. Table 3-5 shows the modified minimum specifications and a revised order of importance to be used in the far-term system evaluation. These specifications are utilized in Table 3-6 for comparing the far term systems.

## Discussion of Parameters

The system parameters being used to evaluate and compare vocoders are listed in Tables 3-1 and 3-5. These parameters are not of equal importance. The relative importance of each is indicated by the parameter weight in Tables 3-2 and 3-6. The order of importance derives from estimated requirements of the Air Force and needs of the LPI Comm ADP. These parameters consistute as complete a set as is possible at this time. Even within this group, quite often some parameters are not available on a system. A discussion of each is given below.

Data Rate. Speech information rate is the most important parameter because it directly affects the vocoder gains towards improved LPI capabilities as discussed in Chapter II and in Appendix A2. Systems are being researched with data rates ranging from 75 b/s to 16000 b/s. The maximum rate of 2400 b/s was

TABLE 3-5

SYSTEM PARAMETERS AND SPECIFICATIONS FOR
FAR TERM ANALYSIS

| Parameter | |
|---|---|
| Data Rate | $\leq$ 800 bits/sec |
| Intelligibility | $\geq$ 88% DRT |
| Size | $\leq$ 200 $in^3$ |
| Weight | $\leq$ 2.5 lb. |
| Power | $\leq$ 5 W |
| Input Probability of Error | $\geq 10^{-3}$ |
| Production Cost | $\leq$ \$1000/unit |
| System Availability | Research into modification of current systems (minimum) |
| Processing Delay | $\leq$ 100 msec |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

TABLE 3-6

FIGURE OF MERIT ANALYSIS SYSTEM FOR LASARS IMPLEMENTATION SYSTEM (Sample Chart)

SYSTEM: ______________________

SOURCE: ______________________

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | | |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | | |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | | |
| WEIGHT (lb) | 0.080 | <0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | | |
| POWER CONSUMPTION (Watts) | 0.080 | <0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | | |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5 \times 10^{-2}$ | — | $10^{-2}$ | — | $5 \times 10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | | |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | >1000 | | |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | | |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | | |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | | |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | | |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | | |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | |

established because the Air Force has already approved a 2400 b/s LPC vocoder for a different implementation. Air Force, Navy, DoD, and NATO specifications exist at this time describing vocoder output formats conforming to a specific form of LPC (LPC-10) at 2400 b/s. These specifications do not constrain advanced development implementations.

Intelligibility. Intelligibility is a key parameter and is nearly as important as data rate. Unlike telephone conversations with nearly unlimited context, aircraft mission communications must achieve a very high transfer rate of information with extremely limited context and without repetition. As the mission criticality increases, the contextual support available decreases. At the present time, as a result of testing convenience, there is only one widely used, quantitative measure of vocoder intelligibility. This is the Diagnostic Rhyme Test (DRT) developed and conducted by Dr. William Voiers of Dynastat, Inc., Austin, Texas. [139-146] Voiers, Air Force personnel and others (10, 68, 100, 119, 123, 135, 145) have determined that systems scoring approximately 80 percent on the DRT provide reasonable acceptability of the intelligibility of the transmitted speech signal. At this time, new, realistic, conversational tests are being developed (113, 125) which could produce new results before the implementation of the LASARS. All intelligibility scores are given for accustically benign environments.

Input Probability of Error. In any communications link, errors occur which distort the incoming data. These errors are the

result of a variety of sources such as lightning, sunspots, atmospheric temperature fluctuations, other communication transmissions, multipath transmission, etc. Because of this, communication systems must be tolerant to a certain number of bit errors. The number of errors the system can tolerate directly affects the signal power level out of the transmitter. Because an LPI communication system is aimed at operation at marginal levels, the more errors the system can tolerate, the larger the signal power reductions can be, making the system more attractive for LPI usage. It is desired that the synthesizer of the brassboard insertion system require for operation an input probability of error, $P_e$, no smaller than $10^{-3}$.

Physical Parameters. Size, weight, and power consumption are highly correlated factors. They are largely technology dependent. Higher levels of circuit integration mean reduced size, reduced weight, and reduced power consumption. A microprocessor analyzer/synthesizer is smaller, lighter, and has a lower power consumption than a filter bank system. The brassboard system tested in the late 1980s is being designed to test the overall LPI system concepts. It will be flown in the cargo bay of a military cargo aircraft on pallet-mounted racks. At this time optimum size, weight, and power constrictions will not apply. The specifications are chosen somewhat arbitrarily to insure the equipment can be physically handled, reasonably easily. It is believed that approximately two cubic feet and fifty pounds should be a limit. The power requirements are also not too restrictive. The test system will have an independent power source available to provide whatever power is needed. One hundred

watts is projected to be the maximum power necessary for a flight testable brassboard system. More restrictive requirements apply to the LASARS implementation because they would be permanently mounted in much smaller aircraft.

Data Input/Output Delay. Processing delay is important in terms two of aspects. If the system does not operate in real-time, it is unacceptable. Real-time systems generate a continuous output given a continuous input without having to pause to perform processing and without overloading the system causing a loss of data. The second aspect is pipeline delay. This is the length of time needed to provide an initial output for an initial input. In two-way conversations, any time lag greater than 100 milliseconds is noticeable and delays of 250 milliseconds or more make conversations hard (almost impossible) to conduct in a strategic or tactical environment where rapid communications are necessary.

Availability. System availability occurs in two phases, near- and far-term. Near-term systems exist as engineering prototypes with production models available in three to five years or as systems currently in production. Far-term systems include the near-term systems, laboratory models available for production in six to ten years, and also simulations of modifications to current laboratory models, engineering prototypes, or production systems. The "far-term-only" systems are not applicable to the brassboard development but any of them on which research is continued are appropriate considerations for the LASARS implementation in the mid-1990s.

Cost. Production cost is the specification utilized in this analysis. It is not the most accurate measure of system cost. Life cycle cost is more comprehensive. It includes development costs, production costs, cost for purchasing and stocking spare parts, maintenance costs, costs for training maintenance personnel, and replacement costs--all based on the life of the system and the life of the host equipment (i.e., the airframe). This information is not available from the vendor and lack of personal experience forbids making estimates.

In the purchase of one or two systems for brassboard testing, cost is relatively unimportant. More can be paid per individual item to test a concept than to implement one. The managers of the LPI program have determined that $40,000 is not too much to pay for a test system. For production purposes, the vocoder portion of the LASARS should cost as little as possible, preferably less than $1,000 each. The production cost figures utilized are estimates only. This figure is usually established through bids and varies with quantity.

Binary Decision Parameters. Speaker dependence, vocabulary dependence and system learning time constitute binary decision parameters. Any system which requires a specific speaker to achieve the required intelligibility or to be voice recognizable by the listener is unacceptable. Vocabulary dependent systems, usually utilizing some type of look-up table, are unacceptable because of the wide range of applications for the vocoder and the wide range of mission requirements within a single application. These parameters, then, establish a "go/no go" limit on the systems under consideration.

Several systems are under investigation which utilize a look-up table to find a "best-match" pattern to the analyzed segment of speech. These systems update the patterns in memory utilizing a "least accessed, first replaced" algorithm. In this type of system, intelligibility is maintained but the ability to recognize the speaker occurs only after the new speaker has caused enough spectral patterns to be replaced. The pattern replacement is called system "training" or "learning" time. Instantaneous speaker recognition is necessary in short duration, high volume communication environments involving many different speakers. Therefore, the "learning" time required must provide "almost instantaneous" updating of the memory patterns. Although intelligibility is maintained, a very slight degradation occurs initially which is returned to normal as the system "learns." This increases the requirements for rapid updating of the system. The systems should require no training time.

Most, if not all, of the specifications of the parameters discussed will be tightened when final consideration is made to determine the 1990's LASARS applicable vocoder. Not much improvement in intelligibility is expected although more robust operation is expected. More immunity to acoustic noise in the analysis environment is a result of the improvements in the robustness. Also included in this is more naturalness (less mechanical sounding) in the synthesized speech. The new parameter specifications are those given in Table 3-5.

# CHAPTER IV

## PRESENTATION OF VOCODER SYSTEMS

### Overview

The research conducted identified more than thirty vocoder systems existing as production equipment, working engineering prototype models, working laboratory models, or as computer software simulations. This chapter presents each system with some general, qualitative information including developing organization, analysis/synthesis method, strengths and weaknesses. A table with the quantitative system parameter values is provided for each system.

### Vocoder System Presentation

Table 4-1 gives the developing organization, system nomenclature, and the references in the Bibliography for each system. As indicated in Chapter III, systems which exist only as computer software simulations are inappropriate to be considered for either phase of the LPI Comm ADP by reason of nonavailability. Table 4-2 lists these nonavailable systems. Additional information on these systems can be found in the references cited in Table 4-1 and will not be included here. Most of the systems listed in Table 4-2 do not run in real-time at the present and several could have been eliminated for

TABLE 4-1

VOCODER SYSTEMS IDENTIFIED

| Developer/Producer | System/Nomenclautre | Reference |
|---|---|---|
| ITT/Tri Tac | CV-3591 (ANDVT) | 20, 39, 61, 63, 64, 129, 131 |
| USAF/ITT | Frame Predictive LPC | 125, 149 |
| TI | VIS-Speech Processor | 72 |
| TI | Time Encoded LPC Roots | 97 |
| Motorola | ATMMRP | 18, 30, 89 |
| Motorola | Manpack | 18, 30, 89 |
| Motorola | MNSVS | 18, 30, 89 |
| E-Systems | CV-333A/U | 9, 13, 22, 24, 147 |
| E-Systems | CV-3333/U | 9, 13, 25, 26, 147 |
| E-Systems | CV-3670/A | 9, 13, 23, 25, 27, 147 |
| E-Systems | LPC-24 | 9, 13, 28, 147 |
| GTE | MRD-2000G | 51, 52, 79, 148 |
| GTE | UVD-2000 | 51, 52, 79, 148 |
| GTE | CV-3832 (MRVT) | 53, 79, 148 |
| GTE | TDHS | 73 |
| MIT Lincoln Labs (LL) | Compact LPC | 29, 57 |
| LL | Adaptive Subband Format Analysis | 87 |
| LL | SEE | 99, 100 |
| LL | 800 b/s SEE | 98, 100, 101 |
| LL | Wideband SEE | 100 |
| LL | Frame Fill LPC | 10, 11, 100, 101 |
| LL | Channel | 48 |
| LL | Pattern Matching Channel | 43, 44 |
| LL | Vector Quantized LPC | 10, 100, 101 |
| LL | Homomorphic Prediction | 69 |
| Bolt, Baranek, & Newman (BBN) | Segment Quantization | 108 |
| BBN | Single Frame Quantization | 108 |
| BBN | HDV LPC | 138 |
| BBN | Variable Order Markov | 107 |
| Naval Research Lab (NRL) | Linear Predictive Formant | 67 |
| NRL | Line Spectrum Pairs | 37 |
| NRL/TRW Corp. | Vector Quantized LPC | 36, 37, 66 |
| Stanford Research Inst. | RELP | 133, 134 |

TABLE 4-1--Continued

| | | |
|---|---|---|
| Univ. of Notre Dame | RELP | 17 |
| Signal Technology, Inc. | Vector Quantized LPC | 150, 151 |
| Korean Advanced Inst. of Science | Low-Rate Digital Formant | 132 |
| DoD | Differential LPC | 38 |
| DoD/CNR, Inc. | LPC-10 Formant | 92 |

TABLE 4-2

NONAVAILABLE SYSTEMS OR ALGORITHMS

| Developer/Producer | System |
|---|---|
| GTE | TDHS |
| Stanford Research Inst. | RELP |
| Univ. of Notre Dame | RELP |
| Korean Advanced Inst. of Science | Low-Rate Digital Formant |
| LL | Pattern Matching Channel |
| LL | Adaptive Subband Formant Analysis |
| LL | Wideband SEE |
| LL | SEE |
| LL | 800 b/s SEE |
| LL | Vector Quantized LPC |
| LL | Homomorphic Prediction |
| TI | Time Encoded LPC Roots |
| BBN | Segment Quantization |
| BBN | Single FramE Quantization |
| BBN | HDV LPC |
| BBN | Variable Order Markov |
| NRL | Linear Predictive Formant |
| NRL | Line Spectrum Pairs |
| DoD | Differential LPC |
| DoD/CNR, Inc. | LPC-10 Formant |
| Signal Technology, Inc. | Vector Quantized LPC |

other reasons such as data rate too high, single speaker restrictions, or speech intelligibility too low.

Twenty one systems or algorithms have been eliminated by being nonavailable in the desired time frame. This leaves seventeen systems to be considered in either the near-term or far-term application evaluations. Table 4-3 lists the systems appropriate for the brassboard insertion as determined by the availability as given in Chapter III. Table 4-4 lists the systems which, in addition to those in Table 4-3, are appropriate for inclusion in the production LPI Comm radio system.

TABLE 4-3

SYSTEMS CONSIDERED FOR NEAR-TERM BRASSBOARD

| Developer/Producer | System |
|---|---|
| ITT | CV-3591 (ANDVT) |
| Motorola | ATMMRP |
| Motorola | Manpack (83-2791) |
| Motorola | Miniturized NSV System |
| E-Systems | CV-3333 A/L |
| E-Systems | CV-3333/U |
| E-Systems | CV-3670/A |
| E-Systems | LPC-24 |
| GTE | MRD-2000G |
| GTE | UVD-2000 |
| GTE | CV-3832 (MRVT) |
| NRL/TRW Corp. | Vector Quantized LPC |

<u>Vocoder System Descriptions--Brasboard Applicable</u>

In this section the vocoder systems listed in Table 4-3 are described. These are the systems applicable for the late 1980s brassboard insertion. The descriptions consist of a paragraph giving the developer/producer, engineering status, information about the

analysis/synthesis methods used, and any appropriate qualitative comments for each system. This is followed by a table listing all of the quantitative data available for each system.

TABLE 4-4

ADDITIONAL SYSTEMS FOR FAR-TERM PRODUCTION CONSIDERATIONS

| Developer/Producer | System |
|---|---|
| USAF/ITT | Frame Predictive LPC |
| LL | Compact LPC |
| LL | Frame Fill LPC |
| LL | Channel |
| TI | VIS-Speech Processor |

<u>ITT--CV 3591 (ANDVT) Vocoder</u>. The Advanced Narrowband Digital Voice Terminal (ANDVT) is a DoD approved, 2400 bit/second, LPC vocoder. This government-wide system is the result of extensive research by ITT with a tri-service research committee. It is scheduled to go into production within a couple of months with all service branches purchasing units. The system contains an adaptive acoustical noise cancellation algorithm for improving speech intelligibility in tactical environments. It has two transmission modes, HF and Line-of-Sight. This unit is to be utilized in all Air Force communication systems now existing which require low data rates for secure voice applications. The format of this system has even been accepted as a North Atlantic Treaty Organization (NATO) standard. The coding consists of a mixture of pitch and amplitude semi-logarithmic coefficients, log-area-ratio and linear coefficients for the filter

transfer function specifications, and error detection and correction bits. The system also has the capability to transmit nonvoice data at 300, 600, 1200, and 2400 b/s. Table 4-5 lists the quantitative parameter values.

TABLE 4-5

ITT CV-3591

| Parameter | Value |
|---|---|
| Vocoder Method | LPC-10 |
| Data Rate | 2400 |
| Intelligibility | Approximately 83% |
| Input $P_e$ | $10^{-2}$ |
| Size | 750 $in^3$ |
| Weight | 20 lb. |
| Power | 45 W |
| Processing Delay | 50-60 mS |
| Availability | Production |
| Production Cost | $20,000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

Motorola--ATMMRP Vocoder. The Advanced Technology Model Multi Rate Processor (ATMMRP) LPC Vocoder is a small, low power, 2400 b/s, full duplex, LPC vocoder. It is compatible with the DoD ANDVT. It employs an MC68000 microcomputer and CMOS circuits. It utilizes partial correlation (PARCOR) analysis to determine the filter transfer function specifications. An additional output is provided for Residual Excited LPC (RELP) coding (9600 b/s). The system consists of several microprogrammed digital signal processing ICs. It currently exists at Motorola, Inc., Scottsdale, Arizona, as an

operational engineering prototype model. Table 4-6 lists the quantitative values used in the comparison.

TABLE 4-6

MOTORLA ATMMRP

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400 |
| Intelligibility | Approximately 89% |
| Input $P_e$ | $5x10^{-3}$ |
| Size | 440 $in^3$ |
| Weight | 9.1 lb. |
| Power | 4.2 W |
| Processing Delay | 50-60 |
| Availability | Working Model |
| Production Cost | $6000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

Motorola--Manpack (83-2791) Vocoder. The Manpack is an extension of the ATMMRP chip set with further reductions in size. It implements parallel processing techniques in three signal processing chips to accomplish the size and power reductions. The RELP output is deleted but compatibility with the ANDVT is maintained. The V/UV decision and pitch tracking algorithms have been improved to optimize operation performance for noisy, rapid communications. A high performance automatic gain control has been designed and included especially for the rapid communication environment. The system exists as a working, engineering prototype model. Table 4-7 lists the quantitative parameter values.

TABLE 4-7

MOTOROLA MANPACK

| Parameter | Value |
|---|---|
| Vocoder Method | LPC-10 |
| Data Rate | 2400 |
| Intelligibility | Approximately 89% |
| Input $P_e$ | $5 \times 10^{-3}$ |
| Size | 90 $in^3$ |
| Weight | 4.1 lb. |
| Power | 2 W |
| Processing Delay | 50-60 mS |
| Availability | Working Model |
| Production Cost | $7000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

Motorola--MNSVS Vocoder. The Miniaturized Narrowband Secure Voice System (MNSVS) is a further reduction in size over the Manpack. It utilizes flatpack and leaded chip carrier technology over dual inline packages (DIP) to accomplish these reductions. Included in the system is a dedicated microprocessor to provide a variety of security levels. This system also exists as a working model at Motorola, Inc., Scottsdale, Arizona. Table 4-8 lists the quantitative values. The Motorola, Inc., Communications Division Product Information Report [89] contains photographs of all three of Motorola's vocoders presented here.

E-Systems--CV-3333/U. The CV-3333/U Mil Spec Digital Speech Processor is a full/half duplex, 2400 b/s, channel vocoder being produced for the U.S. Navy. Its output can be multiplexed with

other data streams to allow simultaneous voice and data transmission. The system can be operated from a standard telephone input. It is compatible with the HY-2 channel vocoder which it is replacing. (The HY-2 is constructed with discrete components (119)). Table 4.9 gives the parameter values used in the comparisons.

TABLE 4-8

MOTOROLA MNSVS

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400 |
| Intelligibility | Approximately 89% |
| Input $P_e$ | $10^{-2}$ |
| Size | 20 $in^3$ |
| Weight | 1.5 lb. |
| Power | 2 W |
| Processing Delay | 50-60 mS |
| Availability | Working Model |
| Production Cost | $8000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

E-Systems--CV-3333 A/U. The CV-3333 A/U Audio-Digital Converter is a full/half duplex, 2400 b/s, LPC vocoder. The system is in production at E-Systems Garland Division. It is compatible with the ANDVT. Also provided is compatibility with the HY-2 channel vocoder. The LPC filter parameters are specified through the use of reflection coefficients with standard V/UV decision and pitch tracking algorithms employed. Like the CV-3333/u, it can be

multiplexed with other data streams. Both systems contain self-test subroutines for quick fault isolation and repair. Repair is accomplished by board replacement. Table 4-10 gives the quantitative values.

TABLE 4-9

E-SYSTEMS CV-3333/U

| Parameter | Value |
|---|---|
| Vocoder Method | Channel |
| Data Rate | 2400 |
| Intelligibility | Approximately 88% |
| Input $P_e$ | $2 \times 10^{-3}$ |
| Size | 2800 $in^3$ |
| Weight | 55 lb. |
| Power | 200 W |
| Processing Delay | Approximately 75 mS |
| Availability | Production |
| Production Cost | Approximately $25,000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

E-Systems--CV-3670/A. The CV-3670/A Airborne Digital Speech Processor is a 2400 b/s, LPC vocoder/secure voice system smaller than the CV-3333 A/U (half the parts count). It has improved intelligibility and quality. The CV-3670/A is currently in use aboard the AWACS and other USAF aircraft. The system is ANDVT compatible. The speech signal analysis is performed using an E-Systems proprietary algorithm utilizing the reflection coefficients. It is compatible with standard security equipment and has been designed for minimum electromagnetic emissions. It interconnects for I/O with the aircraft intercom system. The basic unit can be mounted almost anywhere

TABLE 4-10

E-SYSTEMS CV-3333 A/U

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400 |
| Intelligibility | Approximately 90% |
| Input $P_e$ | $2 \times 10^{-3}$ |
| Size | 2800 $in^3$ |
| Weight | 45 lb. |
| Power | 100 W |
| Processing Delay | Approximately 75 mS |
| Availability | Production |
| Production Cost | Approximately $22,500 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

through use of a remote control unit provided. As indicated, this unit exists in the Air Force inventory. Table 4-11 gives the quantitative data. The CV-3333, CV-3333 A/U, CV-3670/A, all interface with MIL STD-188 cryptography units.

E-Systems--Model LPC-24. The Model LPC-24 Digital Speech Processor is a 2400 b/s LPC commercial vocoder which has been sold in quantity internationally. It has high speech quality. It is designed for operation on a dedicated network. Optional equipment is available to allow the LPC-24 to be operated as a dedicated single use terminal and to expand the LPC-24 capability to include a teleprinter channel. Table 4-12 lists the system quantitative parameter values.

TABLE 4-11

E-SYSTEMS CV-3670/A
(remote unit/remote control)

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400 |
| Intelligibility | Approximately 90% |
| Input $P_e$ | $2 \times 10^{-3}$ |
| Size | 728 $in^3$ / 78 $in^3$ |
| Weight | 20 lb. / 1 lb. |
| Power | 90 W / 7 W |
| Processing Delay | Approximately 75 mS |
| Availability | Production |
| Production Cost | Approximately $50,000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

TABLE 4-12

E-SYSTEMS LPC-24

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400 |
| Intelligibility | Approximately 90% |
| Input $P_e$ | $2 \times 10^{-3}$ |
| Size | 390 $in^3$ |
| Weight | 20 lb. |
| Power | 100 W |
| Processing Delay | Approximately 75 mS |
| Availability | Production |
| Production Cost | Approximately $10,000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

<u>GTE Systems--MRD-2000G</u>. The MRD-2000G Voice Digitizers is a 2400 b/s, LPC vocoder designed for military, narrowband secure voice systems. It utilizes a GTE proprietary implementation of LPC analysis/synthesis designated LPC 10/42. The LPC output is ANDVT compatible. The system can be supplied with optional, switch selectable voice processing algorithms. These are adaptive predictive coding (APC) and sub-band coding (SBC) at 7200 b/s and 9600 b/s respectively. A second option, which is added to the basic LPC to make the system more compatible with military systems, is a set of channel vocoder outputs. These outputs are compatible with E-System's CV-3333 vocoders, the HY-2, or Great Britian's Belgarde channel vocoder (another discrete component system). The system contains echo suppression circuitry with 60 dB of echo suppression and self-test circuitry for fast fault isolation. Military Standard MIL STD-188C provides digital interface to external encryption devices and other data communications equipment. I/O is through a handset telephone receiver. Table 4-13 lists the quantitative parameter values.

<u>GTE Systems--UVD-2000</u>. The UVD-2000 Voice Digitizer is a commercial version of the MRD-2000G. It utilizes a GTE proprietary LPC-10 algorithm (10 pole model as in the ANDVT) for the analysis/synthesis system. All of the options and included features of the MRD-2000G are also available. A standard RS-232C interface provides interfacing for the A/D and D/A circuits. Table 4-14 gives the quantitative values.

TABLE 4-13

GTE SYSTEMS MRD 2000G

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400 |
| Intelligibility | Approximately 90% |
| Input $P_e$ | $10^{-3}$ |
| Size | 1666 $in^3$ |
| Weight | 30 lb. |
| Power | 70 W |
| Processing Delay | 50-60 mS |
| Availability | Production |
| Production Cost | Approximately $15,000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

TABLE 4-14

GTE SYSTEMS UVD-2000

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400 |
| Intelligibility | Approximately 90% |
| Input $P_e$ | $10^{-3}$ |
| Size | 1020 $in^3$ |
| Weight | 20 lb. |
| Power | 60 W |
| Processing Delay | 50-60 mS |
| Availability | Production |
| Production Cost | Approximately $10,000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

GTE Systems--CV-3832 MRVT. The GTE Systems Multiple Rate Voice Terminal (MRVT) is an LPC-based multiple rate vocoder. It provides simultaneous outputs of 16000, 9600, and 2400 b/s. Each output is coded independently. The 2400 b/s output forms the basis with data bits added to the stream to generate the higher data rates. This forms an imbedded data scheme. This modem can converse with vocoders at different data rates simultaneously. It also can provide interface between two systems at different rates with the quality limited to that of the system with the lowest data rate. The 9600 b/s output is of the RELP formant (prediction residual bits are utilized, see Table 4-1 for more information and references). The 16000 b/s stream is approximately telephone toll quality. Built in test capabilities are also provided. The 2400 b/s data stream is ANDVT compatible. Table 4-15 lists the system parameter values.

NRL/TRW Corp.--Vector Quantized LCP Vocoder. The Naval Research Laboratory (NRL) is currently having TRW build a Vector Quantized LPC Low Data Rate Voice Terminal (LDRVT) engineering prototype. The system has two switch selectable outputs, standard 2400 b/s LPC and vector quantized LPC at 800 b/s. Vector quantizing occurs by matching the reflection coefficients of each data frame to a set of stored patterns in memory and then transmitting the pattern index instead of the coefficients. The 2400 b/s stream is ANDVT compatible. This system provides the ability to provide an interface between systems at each data rate with quality limited to that of the lower rate system. Some quality and intelligibility degradation occurs because any reasonably sized set of patterns cannot completely model

TABLE 4-15

GTE SYSTEMS MRVT

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400/9600/16000 |
| Intelligibility | Approximately 90% at 2400 b/s |
| Input $P_e$ | $10^{-3}$ |
| Size | 3242 $in^3$ |
| Weight | 55 lb. |
| Power | 200 W |
| Processing Delay | Approximately 60 mS |
| Availability | Production |
| Production Cost | Approximately $25,000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

the full range of speech production. Table 4-16 gives the quantitative values used for comparison purposes.

Vocoder System Descriptions--LASARS Applicable

The preceding paragraphs described those systems which exist as some sort of functioning hardware. These systems along with those listed in Table 4-4 are considered to be a reasonably complete list of systems which should be through all testing stages and be production implementable by approximately 1992 or 1993 for inclusion in the LASARS. This section describes the systems listed in Table 4-4.

USAF/ITT--Frame Predictive LPC Vocoder. Through an Air Force monitored contract with ITT, a 400 b/s vocoder is being designed. This vocoder implements 2400 b/s LPC which is then vector quantized to

TABLE 4-16

NRL/TRW VECTOR QUANTIZED LPC

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 800 |
| Intelligibility | 84% |
| Input $P_e$ | (not tested) |
| Size | 2400 $in^3$ |
| Weight | 30 lb. |
| Power | 90 W |
| Processing Delay | 250 mS |
| Availability | Engineering Prototype |
| Production Cost | Not Available |
| Speaker Depend. | None |
| Vocabulary Depend. | None |
| System "Learning" Time | None |

800 b/s. The 800 b/s data stream is then frame predicted with the pitch, V/UV, and gain parameters coded through fake process trellis coding utilizing variable rate coding to further reduce the data rate to under 400 b/s (149). Frame prediction techniques are employed to remove the frame-to-frame redundancy in the LPC filter parameters through the use of frame repeat coding. Given a frame of data transmitted, if the following frame is "close enough" using some distortion measure as in Itakura [50] or Wong [149] it is not transmitted. Instead, a one-bit/frame repetition flag (repeat/not repeat) is transmitted. The fake process trellis coding of the excitation parameters is performed independently of the vector quantization and the frame prediction. A search algorithm, e.g., Viterbi or ML (149), is employed to determine the code to minimize the

expected distance between the input excitation parameters and the encoded output.

This frame predictive LPC method currently exists as a software algorithm to modify an LPC input and is hosted on a VAX computer. All of the additional processing is implementable with programmable, signal processing chips which could be added to ITT's existing ANDVT design with little re-engineering required. If given the "go-ahead" (funds) ITT states they could have a working model in less than a year (125). Table 4-17 lists the parameter specifications.

TABLE 4-17

USAF/ITT FRAME PREDICTIVE

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 400 |
| Intelligibility | 78.9% |
| Input $P_e$ | $10^{-2}$ |
| Size | (Unknown |
| Weight | Mainframe |
| Power | Simulation) |
| Processing Delay | Approximately 300 mS |
| Availability | Simulation |
| Production Cost | Not Available |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

<u>LL--Compact LPC Vocoder</u>. MIT's LL located at Hanscom AFB, Massachusetts, is performing extensive research into vocoder design and algorithm improvement. The Compact LPC vocoder is a single card, laboratory model system. It is small, low power, and relatively

inexpensive. It is a 2400 b/s system utilizing only commercially available devices. An autocorrelation analysis is performed to generate the reflection coefficients. An Intel 8085 is utilized to control and supervise the functions within the LPC analyzer, synthesizer, and Gold pitch detector (41). The system is designed for use with a compact packet voice terminal. The system parameters are listed in Table 4-18.

TABLE 4-18

LL COMPACT LPC VOCODER

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400 |
| Intelligibility | 89% |
| Input $P_e$ | Not tested |
| Size | 18 $in^3$ (50-100 $in^3$ with packaging |
| Weight | Approximately .75 lb. |
| Power | 5.5 W |
| Processing Delay | Approximately 90 $m^S$ |
| Availability | Laboratory Model |
| Production Cost | Approximately $1000 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

<u>LL--Channel Vocoder</u>. Most channel vocoder research in the U.S. was dropped with the advent of LPC analysis/synthesis techniques. Since about 1980 there has been an increase in interest in the channel vocoding method. B. Gold (44) at LL has more-or-less led the way in this renewed interest. This LSI design vocoder system, utilizing charge coupled devices (CCDs) , is a switch selectable multi-rate

system with data rates of 1200, 2400, 3600 and 4800 b/s. Quality and intelligibility improvements occur with each increase in data rate. Control and coordination is provided through the use of an Intel 8085A-2 with an additional 8085A-2 performing the Gold pitch extraction and V/UV decision making. The input spectrum is divided into 19 channels for analysis. The coefficients generated specify the filter response of the receiving synthesizer. Input is from a standard telephone handset This system is small, light, and power efficient. LL presently has a working laboratory model. Quantitative values are listed in Table 4-19.

TABLE 4-19

LL CHANNEL VOCODER

| Parameter | Value |
|---|---|
| Vocoder Method | Channel |
| Data Rate | 1.2/2.4/3.6/4.8 kb/s |
| Intelligibility | Not Available |
| Input $P_e$ | Not Available |
| Size | 215 $in^3$ |
| Weight | 7 lb. |
| Power | 5.3 W |
| Processing Delay | Not Available |
| Availability | Laboratory Model |
| Production Cost | Not Available |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

<u>LL--Frame Fill LPC Vocodoer</u>. This LL vocoder is another laboratory model system. It starts with a basic 2400 b/s vocoder and produces a 1200 b/s and a 2400 b/s switchable output. The frame fill

technique is accomplished by deleting every other analysis frame from transmission. Two additional bits must be included to specify to the synthesizer the index of the method to be used to determine how to fill in for the missing frame. Three possibilities exist for regenerating the missing frame, either adjacent frame can be used or some weighted combination of the two transmitted frames can be used. With an input probability of error less than $10^{-3}$ the resynthesized speech is virtually indistinguishable from the original 2400 b/s LPC speech. The quantitative data available is listed in Table 4-20. Reports on this system are as yet unpublished.

TABLE 4-20

LL FRAME FILL LPC

| Parameter | Value |
|---|---|
| Vocoder Method | Frame Fill LPC |
| Data Rate | 1200/2400 |
| Intelligibility | 84% |
| Input $P_e$ | $10^{-2}$ |
| Size | 700 $in^3$ |
| Weight | 22 lb. |
| Power | 70 W |
| Processing Delay | 250 mS |
| Availability | Laboratory Model |
| Production Cost | Not Available |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

<u>TI--VIS-Speech Processor Board</u>. Texas Instruments (TI) has performed extensive speech processing research. Their efforts resulted in some of the earliest speech synthesis systems, most

notably the "Speak and Spell" educational toy line. The Voice Interactive Set (VIS) Speech Processor Board incorporates several functions. It performs LPC analysis to provide limited speaker-dependent speech recognition, limited speaker verification, and vocoding. The board is designed for interfacing with a computer system but I/O, A/D conversion, and D/A conversion can be provided by the inclusion of a Codex with interconnections to almost any analog input/output device (telephone handset, intercom microphone/speaker, etc.). The output format of the reflection coefficients is ANDVT compatible. Good quality results have been obtained (72) with the vocoder operating with an acoustic background of 104 dB (helicopter environment) up to 116 dB (other aircraft environments). This board makes use of advanced device packaging technology to achieve an extremely small, 2400 b/s system. All three functions exist in about 15 square inches of circuit board space. This vocoder is currently available only as an engineering prototype system. The quantitative values are listed in Table 4-21.

In the preceding descriptions ANDVT LPC vocoder and HY-2 channel vocoder compatibility has been stressed, where applicable, because the HY-2 or modifications of it (KY-537, an SSI implementation) are currently in use by the AF and because the ANDVT is being purchased for large scale deployment within the Air Force in the immediate future. Compatibility with either of these systems is not a requirement for the brassboard test system. At the present time, it is not a requirement for the LASARS because it has not been considered (54). This could be changed when final specifications are established for the vocoder to be inserted into the LASARS. Conversations with

several engineers at LL and the Air Force Rome Air Development Center (RADC) pointed out that 400 and 800 b/s LPC vocoders whose 2400 b/s basic structure is ANDVT compatible could also be ANDVT compatible with reduced intelligibility and quality.

TABLE 4-21

TI VIS-SPEECH PROCESSOR PARAMTER VALUES

| Parameter | Value |
|---|---|
| Vocoder Method | LPC |
| Data Rate | 2400 |
| Intelligibility | 88.4% |
| Input $P_e$ | $10^{-5}$ |
| Size | 15 $in^3$ (50-100 $in^3$ with packaging) |
| Weight | 6 oz (0.375 lb.) |
| Power | 15W |
| Processing Delay | 50-60 mS |
| Availability | Engineering Prototype |
| Production Cost | $3,500 |
| Speaker Dependence | None |
| Vocabulary Dependence | None |
| System "Learning" Time | None |

The systems presented are far from the final word in vocoder technology. At this time, totally different concepts are being viewed to provide new methods of performing speech analysis. All of the current schemes are based on the human speech production system. Flanagan [32] and Gold [47] have proposed vocoder analysis methods based upon the properties of the human auditory system. This is thought to be a viable approach because of "the fact that the human peripheral auditory system is a superior signal processor to that of the vocoder" (47). This research is concentrating on duplicating the

functions of the auditory system in electrical hardware. This includes the functions of the outer ear, the inner ear (the hammer, anvil, and stirrup), the coder, the cilia which form the chain which transforms the signal from sound waves to electrical impulses. The Lincoln Laboratories report by Gold and Tierney [47] is an extensive discussion of this concept. This method of vocoding is a very long way from implementation but it should provide excellent results by the end of the century for low-rate, highly intelligible speech transmission.

## CHAPTER V

## OPTIMUM VOCODER SYSTEM SELECTION

### Overview

The chapters preceding this one have provided general descriptions. They have laid the groundwork for the final selections to be made. This chapter applies the methodology developed in Chapter III to each of the systems presented in Chapter IV except for the nonavailable systems. The Figure-of-Merit, $F_s$, is computed individually for each system. The results of these computations are used to identify the optimum systems as candidates for the brassboard effort and the LASARS effort.

### System Evaluations

The evaluation is performed in three stages. A table similar to Table 3-2 is filled out for each system listed in Table 4-3 and described in Chapter IV. The three systems from this group achieving the highest $F_s$ are the most likely candidate systems for the brassboard application. Next, a table similar to Table 3-6 is filled out for each system listed in Tables 4-3 and 4-4 with the additional descriptions also found in Chapter IV. The three systems now achieving the highest $F_s$ are the most likely candidate systems for the LASARS application. Again, those systems listed in Table 4-2 are

deleted from consideration. Finally, the top three candidates in each category are compared qualitatively on the basis of information and features which do not lend themselves to quantitative analysis.

Selection of Brassboard Applicable Systems

The $F_s$ for each applicable system is computed in this section. Tables 5-1(a) through 5-1(l) show these computations. Table 5-2 tabulates these values and shows which systems use estimated values, which systems are renormalized for absent data, and which systems are nonacceptable with the reasons for nonacceptability. As shown in Table 5-2 the top three candidates for the brassboard insertion are the (1) MNSVS, (2) Manpack, and (3) ATMMRP vocoders, all from Motorola. All three of these systems are 2400 b/s, LPC vocoders. They are all ANDVT compatible, therefore, they will interface with the systems the Air Force is currently procuring for their low data rate, secure voice communication systems.

When comparing these vocoders, the only major differences are in the size, weight, and power requirements. These vary for two basic reasons. They are designed for different applications and employ different chip fabrication technologies. The ATMMRP is a desk-top unit similar in appearance to a standard "call director telephone." (89) This vocoder's intended purpose is for use in a fixed-location, secure-voice network. It also interfaces with the Executive Secure Voice Network (ESVN). The Manpack employs an extension of the ATMMRP chip set designed to be a portable unit for use in secure voice radio communication systems. The MNSVS is a single, handheld unit similar

TABLE 5-1 (a)

NEAR-TERM FIGURE OF MERIT; CV-3591 VOCODER

SYSTEM: CV-3591 Advanced Narrowband Digital Voice Terminal (ANDVT)

SOURCE: ITT Defense Communications Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | > 95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 3 | 0.675 |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 5 | 1.000 |
| SIZE ($in^3$) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 4 | 0.320 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 4 | 0.320 |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 3 | 0.120 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 4 | 0.140E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 10 | 0.350 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 3 | 0.060E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 3.585 |

TABLE 5-1 (b)

ATMMRP VOCODER

SYSTEM: Advanced Technology Model Multi Rate Processor LPC Vocoder (ATMMRP)

SOURCE: Motorola Inc., Secure Communications Office

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | >95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 6 | 1.350E |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 3 | 0.600 |
| SIZE ($in^3$) | .080 | <20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 4 | 0.320 |
| WEIGHT (lb) | .080 | <.5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 5 | 0.400 |
| POWER CONSUMPTION (Watts) | .040 | <1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 8 | 0.320 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 4 | 0.140E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 5 | 0.175 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 5 | 0.100E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |

OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL 4.005

TABLE 5-1 (c)

MANPACK VOCODER

SYSTEM: Manpack Portable LPC-10 Vocoder

SOURCE: Motorola Inc., Secure Communications Office

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | > 95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 6 | 1.350E |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 3 | 0.600 |
| SIZE ($in^3$) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 6 | 0.480 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 6 | 0.480 |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 9 | 0.360 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 4 | 0.140E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 5 | 0.175 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 5 | 0.100E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 4.285 |

TABLE 5-1 (d)

MNSVS VOCODER

SYSTEM: Minaturized Narrowband Secure Voice System (MNSVS)

SOURCE: Motorola Inc., Secure Communications Office

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | > 95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 6 | 1.350E |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 5 | 1.000 |
| SIZE ($in^3$) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 10 | 0.800 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 7 | 0.560 |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 9 | 0.360 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 4 | 0.140E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 5 | 0.175 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 5 | 0.100E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| | | | | | | | | | | | | | OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | 5.085 |

TABLE 5-1 (e)

CV-333/U VOCODER

SYSTEM: CV-3333/U Digital Channel Speech Processor

SOURCE: E-Systems, Garland Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | >95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 5 | 1.125 |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 2 | 0.400E |
| SIZE ($in^3$) | .080 | <20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 1 | 0.080 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 60 | > 60 | 0 | 0* |
| POWER CONSUMPTION (Watts) | .040 | <1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 0 | 0* |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 2 | 0.070E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 10 | 0.350 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 2 | 0.040E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 2.665 |

TABLE 5-1 (f)

CV-3333A/U VOCODER

SYSTEM: CV-3333A/U Audio-Digital Converter

SOURCE: E-Systems, Garland Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | >95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 7 | 1.575 |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 2 | 0.400E |
| SIZE ($in^3$) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 1 | 0.080 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 1 | 0.080 |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 1 | 0.040 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 2 | 0.070E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 10 | 0.350 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 3 | 0.060E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| | | | | | | | | | | | | | OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | 3.255 |

TABLE 5-1 (g)

CV-3670/A VOCODER

SYSTEM: CV-3670/A Airborne Digital Speech Processor

SOURCE: E-Systems, Garland Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | >95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 7 | 1.575 |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 2 | 0.400E |
| SIZE ($in^3$) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 3 | 0.240 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 4 | 0.320 |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 1 | 0.040 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 2 | 0.070E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 10 | 0.350 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 0 | 0E* |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| | | | | | | | | | | | | | OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | 3.595 |

TABLE 5-1 (h)

LPC-24 VOCODER

SYSTEM: Linear Predictive Coder Model LPC-24 Digital Speech Processor

SOURCE: E-Systems, Garland Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | > 95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 7 | 1.575 |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 2 | 0.400E |
| SIZE ($in^3$) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 3 | 0.240 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 4 | 0.320 |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 1 | 0.040 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 2 | 0.070E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 10 | 0.350 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 5 | 0.100E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| | | | | | | | | | | | | | OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | 3.695 |

TABLE 5-1 (i)

MRD-2000G VOCODER

SYSTEM: Model MRD-2000G Voice Digitizer

SOURCE: GTE Systems, Communication Systems Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | > 95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 7 | 1.575E |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.200E |
| SIZE (in$^3$) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 2 | 0.460 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 3 | 0.240 |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 2 | 0.080 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 4 | 0.140E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 10 | 0.350 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 4 | 0.080E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 3.425 |

TABLE 5-1 (j)

UVD-2000 VOCODER

SYSTEM: Model UVD-2000 Voice Digitizer

SOURCE: GTE Systems, Communication Systems Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | > 95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 7 | 1.575E |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.200E |
| SIZE ($in^3$) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 3 | 0.240 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 4 | 0.320 |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 3 | 0.120 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 4 | 0.140E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 10 | 0.350 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 5 | 0.100E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 3.645 |

TABLE 5-1 (k)

CV-3832 VOCODER

SYSTEM: CV-3832 Multiple Rate Voice Terminal (MRVT)

SOURCE: GTE Systems, Communication Systems Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 1 | 0.250 |
| INTELLIGIBILITY (% DRT) | .225 | > 95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 7 | 1.575E |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.200E |
| SIZE (in³) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 0 | 0* |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 0 | 0* |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 0 | 0* |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 4 | 0.140E |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 10 | 0.350 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | 2 | 0.040E |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 2.905 |

TABLE 5-1 (1)

VECTOR QUANTIZED LPC VOCODER

SYSTEM: 2400 to 800b/s LPC Rate Converter (Vector Quantized LPC)

SOURCE: NRL with TRW Corp

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | .250 | < 150 | 150 | 200 | 250 | 400 | 600 | 800 | 1200 | 1600 | 2400 | >2400 | 4 | 1.000 |
| INTELLIGIBILITY (% DRT) | .225 | > 95 | 93 | 91 | 90 | 89 | 88 | 87 | 85 | 82 | 80 | < 80 | 3 | 0.675 |
| MAXIMUM INPUT Pe | .200 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.200E |
| SIZE (in³) | .080 | < 20 | 30 | 40 | 60 | 100 | 200 | 400 | 800 | 1600 | 3000 | >3000 | 1 | 0.080 |
| WEIGHT (lb) | .080 | < .5 | .5 | 1 | 2 | 5 | 10 | 20 | 30 | 40 | 50 | > 50 | 3 | 0.240 |
| POWER CONSUMPTION (Watts) | .040 | < 1 | 2 | 4 | 6 | 10 | 15 | 30 | 50 | 75 | 100 | >100 | 1 | 0.040 |
| PROCESSING DELAY (m Sec) | .035 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | >100 | 0 | 0* |
| SYSTEM AVAILABILITY | .035 | Prod | — | — | — | — | Eng Proto | — | — | — | — | Other | 5 | 0.175 |
| PRODUCTION COST ($/unit/1000) | .020 | < 1k | 1k | 2k | 3k | 5k | 10k | 15k | 20k | 30k | 40k | >40k | --- | ------ |
| SPEAKER DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| VOCABULARY DEPENDENCE | .015 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.150 |
| SYSTEM "LEARNING" TIME (msec) | .005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| | | | | | | | | | | OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | 2.816R |

TABLE 5-2

NEAR-TERM SYSTEM COMPARISONS

| System | $F_s$ | Codes† | Comments |
|---|---|---|---|
| CV-3591 (ANDVT) (ITT) | 3.585 | E | |
| ATMMRP (Motorola) | 4.005 | E | #3 |
| Manpack (Motorola) | 4.285 | E | #2 |
| MNSVS (Motorola) | 5.085 | E | #1 |
| CV-3333/U (E-Systems) | 2.665 | *,E | |
| CV-3333A/U (E-Systems) | 3.255 | E | |
| CV-3670/A (E-Systems) | 3.595 | *,E | |
| LPC-24 (E-Systems) | 3.695 | E | |
| MRG-2000G (GTE) | 3.425 | E | |
| UVD-2000 (GTE) | 3.645 | E | |
| CV-3832 (MRVT) (GTE) | 2.905 | *,E | |
| Vector Quantized LPC (TRW) | 2.816 | *,E,R | |

†Codes: * = Nonacceptable System
E = Estimated values used
R = A renormalized $F_s$ for missing values

to a "contempra" telephone handset. The size and power reductions are achieved through the use of flatpack and leaded chip carrier technology rather than DIP chips (89).

A more detailed description of these vocoders is given in the following excerpt from the Motorola, Inc., Product Information Report (89).

ADVANCED TECHNOLOGY MODEL MULTI RATE PROCESSOR LPC VOCODER

The Voice Processing Laboratory located at the Motorola Government Electronics Group in Scottsdale, Arizona, has applied low power LSI digital signal processing capabilities in the development of an Advanced Technology Model Multirate Processor (ATMMRP) LPC vocoder. Developed by Motorola for the Naval Electronics System Command, it is compatible with the Advance Narrowband Digital Voice Terminal (ANDVT) or the Executive Secure Voice Network (ESVN).

The ATMMRP Vocoder is a low power, small size, 2400 Bit per second, full duplex, Linear Predictive voice coder. The voice coder requires approximately 2.2 watts while the LED displays and specialized MIL 188 digital outputs require another 2 watts for a total DC power of 4.2 watts. This is the lowest power LPC Vocoder yet reported. The ATM voice coder weighs 4.15 Kg (9.1 lbs.), and occupies 7,200 cc (440 $in^3$), with the foot print of a typical call director telephone, see figure 1-1 [not included in this report].

All high performance digital signal processing performed by the ATM LPC vocoder is done with custom large scale integration. Voicing, serialization of data, sync acquisition and sync maintenance algorithms, and self-test functions are performed in software in an MC68000 microcomputer.

The ATM LPC vocoder utilizes a family of CMOS integrated circuits which feature a low-cost, low-power consumption, small physical size approach to measurement of the speech parameters. The CMOS DSP IC family consists of an LPC analysis IC, and AMDF pitch extraction IC, and an LPC synthesis IC.

Each IC is a microprogrammed digital signal processor. The analysis (transmit) IC's contain internal ROM programming to process speech directly from an A/D converter and generate parameters common to LPC/RELP vocoder algorithms. Coding of resultant output data is generalized for maximum flexibility.

The Linear Predictive Coding (LPC) Analyzer IC performs PARCOR linear predictive analysis of speech for LPC vocoder and speech recognition systems. This consists of estimating and removing cross-correlation between forward and backward traveling waves in a lattice digital model of the vocal tract. The purpose of the LPC analysis chip is to perform all of the computationally intensive calculations for 10-pole LPC analysis and energy measurement on a single intergrated circuit.

Residual speech output for pitch extraction and RELP coding is also provided. A block diagram showing the architecture of the analyzer IC is shown in Figure 1-2 [Figure 5-1 in this report].

The Average Magnitude Difference Function IC is a high performance digital signal processor, programmed to perform the pitch measurement algorithm used in all Department of Defense LPC-10 vocoders, speaker identification systems, and many speech recognition systems. AMDF is a robust algorithm for measuring pitch periods of speech by finding the time delay at which the speech wave form is most repetitive. The time delay which produces a minimum AMDF is the pitch period. The AMDF operates directly on speech from an A/D Converter and outputs results to a host CPU.

By performing the AMDF analysis in a dedicated integrated circuit, the computation rate associated with pitch and voicing analysis is dramatically reduced. Furthermore, the I/O structure of the AMDF chip is designed to minimize interface requirements on the host processor. Even the simplest processor hosts can utilize the computational power of the AMDF IC. Architecture of the AMDF is shown in Figure 1-3 [Figure 5-2 in this report].

The voice synthesizer integrated circuit (IC) is a microprogrammable CMOS digital signal processor programmed to perform linear predictive coding (LPC) voice synthesis. High quality voice synthesis may be used with residual excitation to achieve a high degree of natrualness in residual excited LPC applications. It also contains sufficient circuitry to operate on internal excitation for pitch-excited LPC applications.

The speech synthesizer IC, like the others, is designed to perform all the computationally intensive arithmetic for speech synthesis while minimizing the load on the host processor.

The microprogramming features of this IC allow it to be used for lattice all-pole filters, lattice all-zero filters, general second-order cascaded sections (FORMANT synthesis), or line spectral pair synthesis (LSP). In addition, it can be microprogrammed to perform special function filters such as band-pass or low-pass filters. The architecture of the synthesizer IC is shown in Figure 1-4 [Figure 5-3 in this report].

## A Manpack Portable LPC 10 Vocoder

A manpack portable LPC-10 Vocoder has been developed which makes substantial size and power performance improvements over existing LPC Vocoders, by extending the ATMMRP chip set as shown in Figure 1-5 [not included in this report]. The remaining LPC-10 algorithmic components are partitioned by the data and process flow graphs into meaningful multi-purpose stand alone single chip computers, resulting in a vocoder that uses 3 VLSI, and 3 LSI components. The digital signal processing algorithms are partitioned as follows: LPC Analysis IC, LPC Synthesis IC, AMDF Pitch Extraction IC. The data flow processes are partitioned into microcomputers as follows: Transmit Pitch and Voicing in processor #1, Transmit AGC in processor #2, and Parameter Quantization and Serialization in processor #3; in the receive mode sync acquisition and maintenance and parameter deserialization in processor

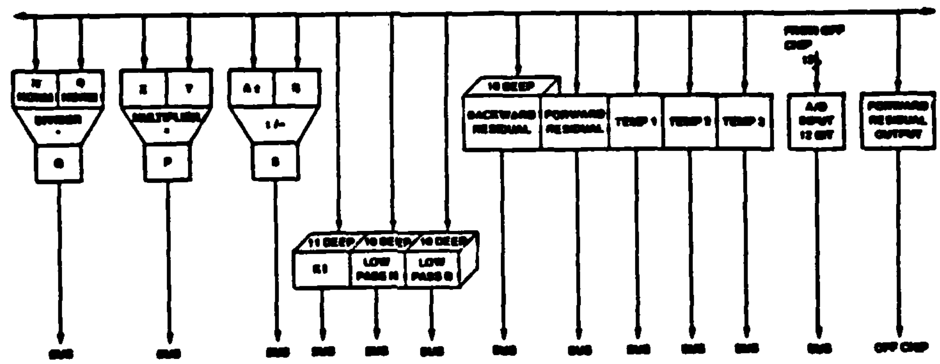


Figure 5-1. Architecture of LPC PARCOR analyzer.
(Taken from ref. 89, p. 4)

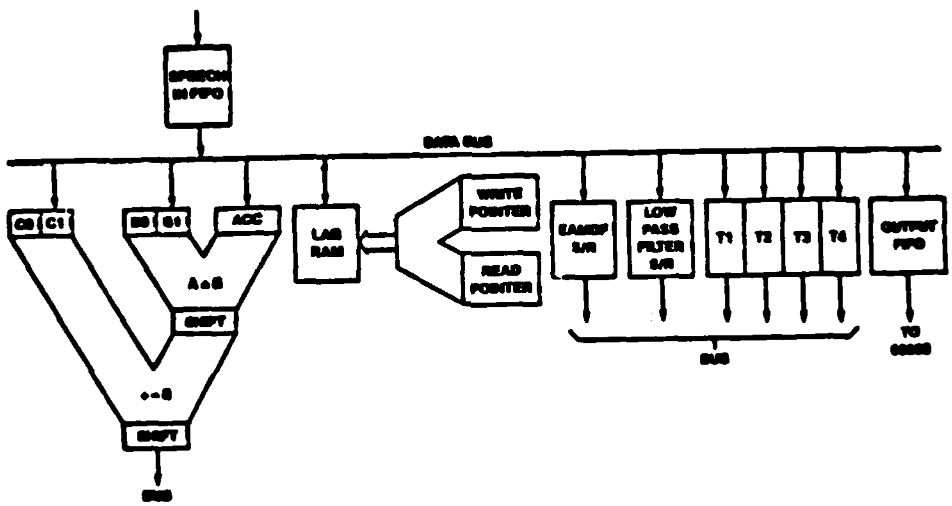


Figure 5-2. Architecture of AMDF IC.
(Taken from ref. 89, p. 4)

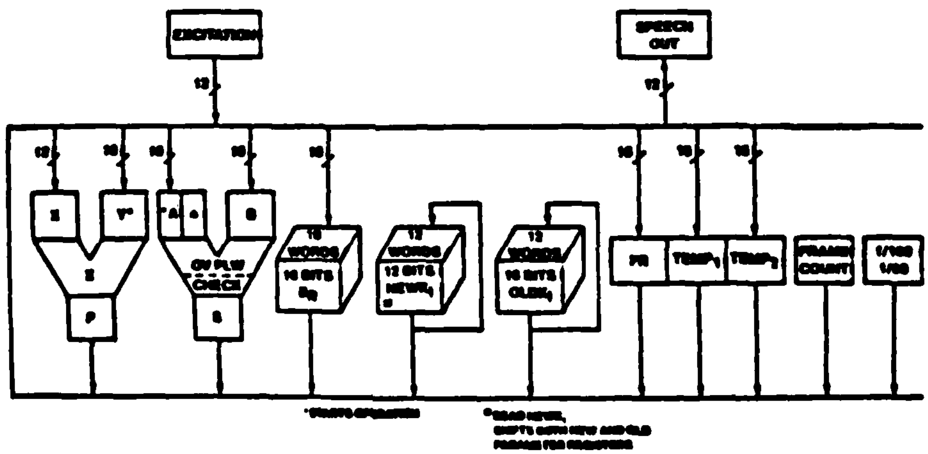


Figure 5-3. Architecture of synthesizer IC.
(Taken from ref. 89, p.4)

#3, error correction and dequantization in processor #2, and interpolation rule implementation in processor #1.

The data flow processor's partitions were greatly affected by the use of single chip computers. The computers have a very limited RAM and ROM space causing the partition to be dependent on program size. The use of single chip computers minimizes external hardware necessary for the vocoder implementation.

The overall block diagram is presented in Figure 1-6 [Figure 5-4 in this report]. The three VLSI speech processing chips, the three microprocessors, and the analog input and output logic comprise the entire half duplex system.

The latest algorithms have been used in Manpack to optimize performance for noisy rapid communication. To accomplish this, the voice/unvoice and pitch tracking algorithms underwent considerable design improvement. Similarly a specially designed high performance automatic gain control has been designed specifically for the rapid communication environment.

Miniaturized Narrowband Secure Voice System

Motorola is now investigating a further miniaturization of the Manpack vocoder by using flatpack and leaded chip carrier technology rather than dual inline packages. The result will be an LPC Vocoder that fits into a "contempra" telephone handset. This further size reduction makes possible an entirely new market for LPC vocoders, due to a small size, portability, and flexibility to be used with a variety of modern technologies. Furthermore, MNSVS also contains a microprocessor based KG controller to enable a variety of secrecy levels of KG to be used with vocoder. The KG control microprocessor mediates link synchronization in non error extending mode with bit error rates up to $10^{-2}$. A photograph of MNSVS is shown in Figure 1-7 [not included in this report].

## Selection of LASARS-Applicable Systems

Tables 5-3(a) and 5-3(q) show the $F_s$ calculations for the far-term applicable systems. These values are tabulated in Table 5-4. As can be seen, none of the systems under consideration achieve all of the currently specified parameter values and all are categorized as nonacceptable for a LASARS implementation.

The Frame Predictive LPC studies of the Air Force with ITT, the MNSVS from Motorola, the Compact LPC Vocoder from Lincoln Laboratories and the TI VIS-Speech Processor ($F_s$ approximately tied with the Compact LPC vocoder) lead the list of systems considered

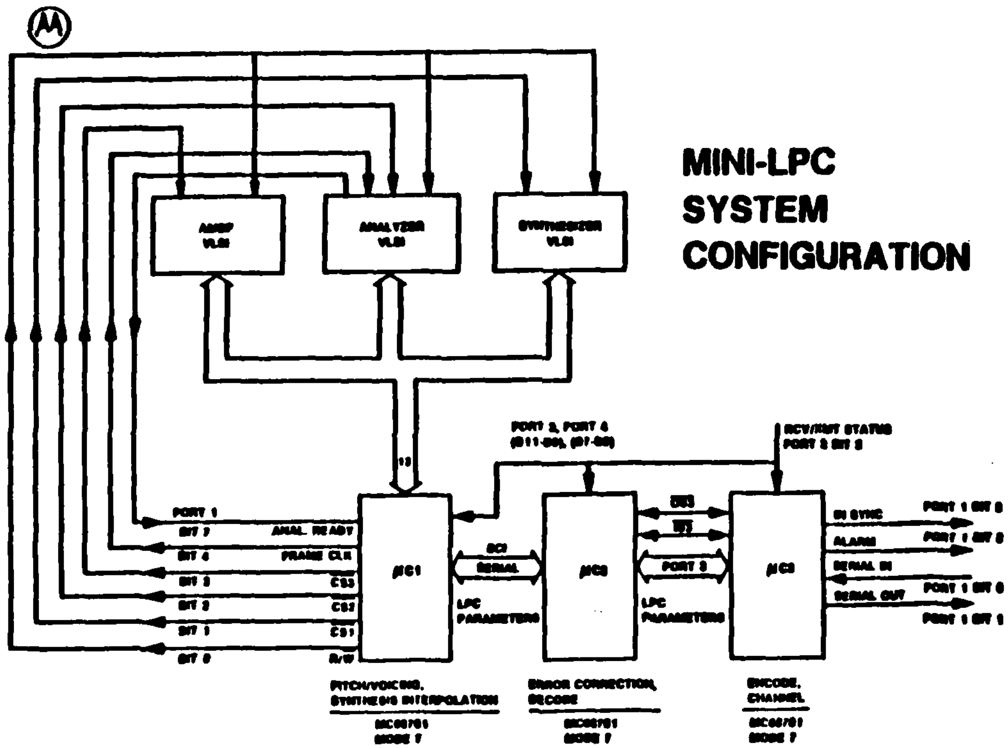



Figure 5-4. Manpack block diagram.
(Taken from ref. 89, p.7)

TABLE 5-3 (a)

FAR-TERM FIGURE OF MERIT; CV-3591 VOCODER

SYSTEM: CV-3591 Advanced Narrowband Digital Voice Terminal (ANDVT)

SOURCE: ITT Defense Communications Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 0 | 0* |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | <0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | <0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 5 | 0.350 |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | >1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 10 | 0.300 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 4 | 0.100E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 0.900 |

TABLE 5-3 (b)

ATMMRP VOCODER

SYSTEM: Advanced Technology Model Mutli Rate Processor LPC Vocoder (ATMMRP)

SOURCE: Motorola Inc., Secure Communications Office

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 2 | 0.500E |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 2 | 0.160 |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 3 | 0.210E |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 7 | 0.210 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 4 | 0.100E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 1.330 |

TABLE 5-3 (c)

MANPACK VOCODER

SYSTEM: Manpack Portable LPC-10 Vocoder

SOURCE: Motorola Inc., Secure Communications Office

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 2 | 0.500E |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 3 | 0.270 |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 6 | 0.480 |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 3 | 0.210 |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 7 | 0.210 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 4 | 0.100E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| | | | | | | | OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | |

TABLE 5-3 (d)

MNSVS VOCODER

**SYSTEM:** Miniaturized Narrowband Secure Voice System (MNSUS)

**SOURCE:** Motorola Inc., Secure Communications Office

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 2 | 0.500E |
| SIZE (in$^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 7 | 0.630 |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 3 | 0.240 |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 6 | 0.480 |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5 \times 10^{-2}$ | — | $10^{-2}$ | — | $5 \times 10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 5 | 0.350 |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 7 | 0.210 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 4 | 0.100E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| | | | | | | | | | | | | | OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | 2.660 |

TABLE 5-3 (e)

CV-3333/U VOCODER

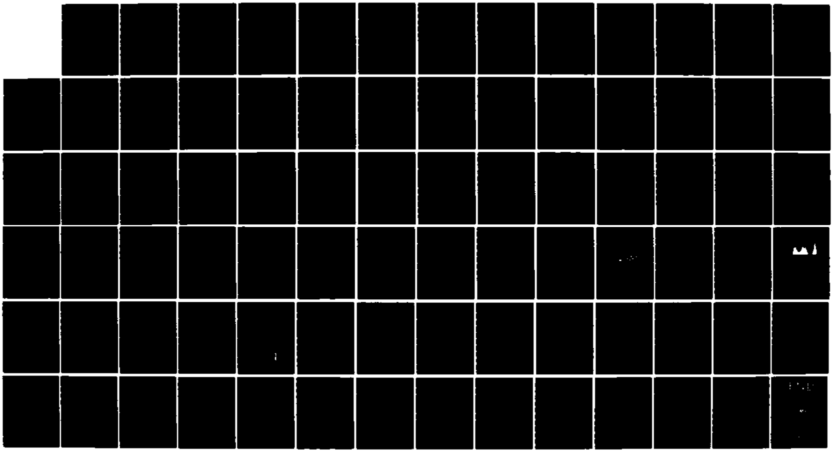
SYSTEM: CV-3333/U Digital Channel Speech Processor

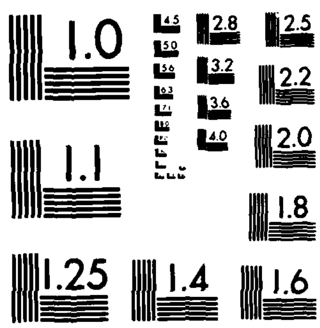
SOURCE: E-Systems, Garland Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 1 | 0.250 |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | <0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | <0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 2 | 0.140 |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 10 | 0.300 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 2 | 0.050E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 0.890 |

AD-A151 919 SURVEY OF NARROW BAND VOCODER TECHNOLOGY(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH W B MCMINN DEC 84 AFIT/CI/NR-85-24T 2/2

UNCLASSIFIED F/G 17/2 NL






MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

TABLE 5-3 (f)

CV-3333 A/U VOCODER

SYSTEM: CV-3333 A/U Audio-Digital Converter

SOURCE: E-System, Garland Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 3 | 0.750 |
| SIZE (in$^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 2 | 0.140 |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 10 | 0.300 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 2 | 0.050E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 1.390 |

TABLE 5-3 (g)

CV-3670/A VOCODER

SYSTEM: CV-3670/A Airborne Digital Speech Processor

SOURCE: E-Systems, Garland Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 3 | 0.750 |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 2 | 0.140 |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 10 | 0.300 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 2 | 0.050E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| | | | | | | | | | | | | | OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | 1.390 |

TABLE 5-3 (h)

LPC-24 VOCODER

SYSTEM: Linear Predictive Coder Model LPC-24 Digital Speech Processor

SOURCE: E-Systems, Garland Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 3 | 0.750 |
| SIZE (in$^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 2 | 0.140E |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 10 | 0.300 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 2 | 0.050E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| | | | | | | | | | | | | OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | 1.390 |

TABLE 5-3 (i)

MRD-2000G VOCODER

SYSTEM: Model MRD-2000G Voice Digitizer

SOURCE: GTE Systems, Communication Systems Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 3 | 0.750E |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.070 |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 10 | 0.300 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 4 | 0.100E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 1.370 |

TABLE 5-3 (j)

UVD-2000 VOCODER

SYSTEM: Model UVD-2000 Voice Digitizer

SOURCE: GTE Systems, Communications Systems Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 3 | 0.750E |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.070E |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 10 | 0.300 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 4 | 0.100E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 1.370 |

TABLE 5-3 (k)

CV-3832 VOCODER

SYSTEM: CV-3832 Multiple Rate Voice Terminal (MRVT)

SOURCE: GTE Systems, Communication Systems Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 3 | 0.750E |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | <0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | <0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5 \times 10^{-2}$ | — | $10^{-2}$ | — | $5 \times 10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.070E |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | >1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 10 | 0.300 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 4 | 0.100E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 1.370 |

TABLE 5-3 (1)

VECTOR QUANTIZED LPC VOCODER

SYSTEM: 2400 to 800 b/s LPC Rate Converter (Vector Quantized LPC)

SOURCE: NRL with TRW Corp.

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 1 | 0.300 |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 0 | 0* |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.070E |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | -- | ------ |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 7 | 0.210 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 0 | 0* |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 0.777R |

TABLE 5-3 (m)

FRAME PREDICTIVE LPC VOCODER

SYSTEM: Vector/Matrix Quantization for Narrow-bandwith Digital Speech Compression

SOURCE: USAF RADC/ITT Defense Communications Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 5 | 1.500 |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 0 | 0* |
| SIZE (in$^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | -- | ------ |
| WEIGHT (lb) | 0.080 | <0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | -- | ------ |
| POWER CONSUMPTION (Watts) | 0.080 | <0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | -- | ------ |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 5 | 0.350 |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | -- | ------ |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 1 | 0.030 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 0 | 0* |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 2.942R |

TABLE 5-3 (n)

COMPACT LPC VOCODER

SYSTEM: Compact, Flexible LPC Vocoder - Commercial Signal Processing Microcomputer

SOURCE: MIT Lincoln Laboratory

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 2 | 0.500 |
| SIZE (in³) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 7 | 0.630 |
| WEIGHT (lb) | 0.080 | <0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 7 | 0.560 |
| POWER CONSUMPTION (Watts) | 0.080 | <0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.070E |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | >1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 4 | 0.120 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 2 | 0.050E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 2.080 |

TABLE 5-3 (a)

CHANNEL VOCODER

SYSTEM: Channel Vocoder

SOURCE: MIT Lincoln Laboratory

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 1 | 0.250E |
| SIZE (in³) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 1 | 0.090 |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 1 | 0.080 |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 1 | 0.070E |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | --- | ------ |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 4 | 0.120 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 2 | 0.050E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 0.862R |

TABLE 5-3 (p)

FRAME FILL LPC

SYSTEM: Frame Fill LPC-10 Vocoder

SOURCE: MIT Lincoln Laboratory

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 0 | 0* |
| SIZE (in$^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 0 | 0* |
| WEIGHT (lb) | 0.080 | < 0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 0 | 0* |
| POWER CONSUMPTION (Watts) | 0.080 | < 0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5\times10^{-2}$ | — | $10^{-2}$ | — | $5\times10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 0 | 0* |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | > 1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 4 | 0.120 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 0 | 0* |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (Fs), TOTAL | | | | | | | | | | | | | | 0.270 |

TABLE 5-3 (q)

VIS-SPEECH PROCESSOR

SYSTEM: Voice Interactive Set (VIS) - Speech Processor

SOURCE: TI Speech Applications Division

| PARAMETER | PARAMETER WEIGHT (Wp) | PARAMETER FIGURE OF MERIT (Fp) MAPPING | | | | | | | | | | | | PARAMETER SCORE (Ps) (Wp x Fp) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Fp | |
| DATA RATE (bits/sec) | 0.300 | < 150 | 150 | 200 | 250 | 300 | 400 | 500 | 600 | 700 | 800 | > 800 | 0 | 0* |
| INTELLIGIBILITY (% DRT) | 0.250 | > 96 | 96 | 95 | 94 | 93 | 92 | 91 | 90 | 89 | 88 | < 88 | 1 | 0.250 |
| SIZE ($in^3$) | 0.090 | < 10 | 10 | 15 | 20 | 25 | 50 | 80 | 100 | 150 | 200 | > 200 | 7 | 0.630 |
| WEIGHT (lb) | 0.080 | <0.5 | 0.5 | 0.75 | 1.0 | 1.25 | 1.5 | 1.75 | 2.0 | 2.25 | 2.5 | > 2.5 | 9 | 0.720 |
| POWER CONSUMPTION (Watts) | 0.080 | <0.5 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 5.0 | > 5 | 0 | 0* |
| MAXIMUM INPUT Pe | 0.070 | $>10^{-1}$ | $10^{-1}$ | — | $5x10^{-2}$ | — | $10^{-2}$ | — | $5x10^{-3}$ | — | $10^{-3}$ | $<10^{-3}$ | 0 | 0* |
| PRODUCTION COST ($/unit/1000) | 0.060 | < 500 | 500 | 550 | 600 | 650 | 700 | 750 | 800 | 900 | 1000 | >1000 | 0 | 0* |
| SYSTEM AVAILABILITY | 0.030 | Prod | — | — | Eng Proto | — | — | Lab Model | — | — | S/W Sim | Other | 7 | 0.210 |
| THROUGHPUT DELAY (m Sec) | 0.025 | < 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 100 | > 100 | 4 | 0.100E |
| SPEAKER DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| VOCABULARY DEPENDENCE | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| SYSTEM "LEARNING" TIME (msec) | 0.005 | NONE | — | — | — | — | — | — | — | — | — | ANY | 10 | 0.050 |
| OVERALL SYSTEM FIGURE OF MERIT (F-), TOTAL | | | | | | | | | | | | | | 2.060 |

TABLE 5-4

FAR-TERM SYSTEM COMPARISONS

| System | $F_s$ | Codes† | Comments |
|---|---|---|---|
| CV-3591 (ANDVT) (ITT) | 0.900 | *,E | |
| ATNNRP (Motorola) | 1.330 | *,E | |
| Manpack (Motorola) | 1.920 | *,E | |
| MNSVS (Motorola) | 2.660 | *,E | #2 |
| CV-3333/U (E-Systems) | 0.890 | * | |
| CV-3333A/U (E-Systems) | 1.390 | * | |
| CV-3670/A (E-Systems) | 1.390 | * | |
| LPC-24 (E-Systems) | 1.390 | *,E | |
| MRG-2000G (GTE) | 1.370 | *,E | |
| UVD-2000 (GTE) | 1.370 | *,E | |
| CV-3832 (MRVT) (GTE) | 1.370 | *,E | |
| Vector Quantized LPC (TRW) | 0.777 | *,E,R | |
| Frame Predictive LPC (ITT) | 2.942 | *,R | #1 |
| Compact LPC | 2.080 | *,E | #3, no card packaging included |
| Channel (LL) | 0.862 | *,E,R | |
| Frame Fill LPC (LL) | 0.270 | *,E | |
| VIS-Speech Processor (TI) | 2.060 | * | Virtually tied with compact LPC, no card packaging included |

†Codes: * = Nonacceptable System
E = Estimated values used
R = A renormalized $F_s$ for missing values

realizable for production and inclusion in a LASARS implementation. None of these combine all of the desirable attributes. The Frame Predictive algorithm along with the NRL/TRW Vector Quantized ($F_s$=.777) system and several other systems (nonavailable) listed in Table 4-2 prove that the desired data rates of 800 b/s and less are possible. The MNSVS, Compact LPC, and the VIS-Speech Processor prove that the required small size necessary in tactical aircraft is also possible. Therefore, with a slight effort at combining the appropriate technologies a small, low-rate vocoder should be possible. The Frame Predictive LPC algorithm will probably require an extra microprocessor or signal processing chip with some additional memory circuits which then could be added to a system such as the MNSVS, Compact LPC, or the VIS-Speech Processor in order to meet the desired specifications.

See the previous excerpt for a discussion on the MNSVS. Additional information on its parameters, characteristics, and chip functions is available from the Motorola Marketing Division, Vicki Crain [18] or Bruce Fette [30]. The Lincoln Laboratory Compact LPC vocoder is presented in detail in Feldman et al. [29], and from Blakenship [10], Gold [44], and Paul [100]. All of the information on the TI VIS-Speech Processor was obtained via the telephone in a private conversation with Langston [72] and is included in Chapter IV.

# CHAPTER VI

# CONCLUSIONS AND RECOMMENDATIONS

## Summary

This research effort has identified, as thoroughly as possible, the current state of technology in vocoder research and production. It has presented an overview of LPI communications and how vocoders form a cornerstone in the LPI conceptual design study. It has described how vocoders operating within the communication link provide significant gains towards the operation of a marginal channel. The speech waveform was discussed in order to provide some insight into the problems vocoder developers have in researching low data rate voice communication methods. It then presented the seven major forms of vocoder algorithms or approaches to speech analysis/synthesis.

In this thesis, a method for quantitatively comparing one system to another was developed with a discussion of each quantitative parameter. Each system of the thirty eight identified was presented in table form. A "first-cut" elimination eliminated all of the nonavailable systems or methods. The remaining systems were individually presented and discussed.

Finally, the comparison method was applied to the available systems in order to identify those <u>most</u> suited for the brassboard

effort to be tested in late 1986 or in 1987 and for the LASARS to be in production and implementable between 1994 and 1996.

Conclusion

At this time several vocoder systems exist either as working models or producton equipment. Of these, three have been identified as the most applicable to the brassboard effort. The objective of this phase of the LPI Comm ADP is to flight test the LPI concepts. The tests will be performed with rack-mounted equipment in the cargo section of an Air Force cargo transport-type aircraft. Considering this, the Manpack or the MNSVS are the most appropriate systems to utilize. The design of the Manpack as shown in Figure 5-3 lends itself to rack mounting. The MNSVS could be attached and then held in a holster when not being used. The ATMMRP would have to have special mounting provided in order to hold it in place.

Currently no vocoder exactly fills the needs of the far-term effort. Several options exist of which one or more will have to be implemented in order to obtain a production model vocoder to fit the LPI needs by 1993. First, the minimum specifications could be reviewed and relaxed so that current models will suffice as production equipment. This would include a detailed analysis of the applications of the LASARS to determine which parameters could or should be modified. Secondly, after the brassboard tests are concluded and if the LPI concept proves viable, a new review of vocoder technology could be conducted with a modified time schedule for LASARS production possible. Finally, additional research funds could be channeled

into the vocoder research efforts with the express purpose of combining the low rate algorithms with the small size technology.

## Recommendations

The recommendations included here are general in nature and are presented as a first consideration for the LPI Comm ADP managers and the LPI Comm Conceptual Design Study contractors. In the brassboard implementation the Motorola Manpack should be used. The only significant difference between it and the higher $F_s$ scoring MNSVS are the size, weight, and power requirements. This system should be chosen because it is rack mountable which makes it more rugged for use in a test environment. Procurement should probably be initiated as soon as possible because the system exists only as an engineering prototype model.

The recommendations for the far-term effort are somewhat harder to make. Additional money should be spent in order to advance the level of vocoder technology. This money should not be immediately dedicated to Air Force contractors currently providing vocoder research to the Air Force. Rather additional organizations such as Motorola, E-Systems, TI, etc. should have an opportunity to bid on this research because they have extensive speech processing research capabilities.

## PUBLICATION ABBREVIATIONS

**Bell Sys Tech J** - The Bell System Technical Journal

**GLOBECOM** - IEEE Global Telecommunications Conference Record

**ICASSP** - IEEE International Conference on Acoustics, Speech, and Signal Processing Record

**ICCR** - IEEE International Communications Conference Record

**IEE Proc** - Proceedings of the Institute of Electronics Engineers

**IEEE Trans ASSP** - IEEE Transactions on Acoustics, Speech, and Signal Processing

**IEEE Trans AU** - IEEE Transactions on Audio and Electroacoustics

**IEEE Trans Comm** - IEEE Transactions on Communications

**J Acoust Soc Am** - Journal of the Acoustical Society of America

**J Audio Eng Soc** - Journal of the Audio Engineering Society

**Proc IEEE** - Proceedings of the IEEE

**Proc Sixth Inter Congr Acoust** - Proceedings of the Sixth International Congress on Acoustics

## BIBLIOGRAPHY


1. Atal, B. S., "New Directions in Speech Coding at Low Bit Rates," GLOBECOM '82, pp. E6.4.1-E6.4.4, (1982).

2. Atal, B. S. and J. R. Remde, "A New Model of LPC Excitation for Producing Natural-Sounding Speech at Lower Bit Rates," ICASSP '82, pp. 614-617, (1982).

3. Atal, B. S. and M. R. Schroeder, "Adaptive Predicive Coding of Speech Signals," Bell Sys Tech J, Vol. 49, pp. 1973-1986, (1970)

4. Atal, B. S. and S. L. Hananer, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave," J Acoust Soc Am, Vol. 50, pp. 637-655, (1971).

5. Bayless, J. W., S. J. Campanella, and A. J. Goldberg, "Voice Signals: Bit-by-Bit," IEEE Spectrum, Vol. 10, No. 10, pp. 28-34, (1973).

6. Beek, B., E. P. Neuberg, and D. C. Hodge, "An Assessment of the Technology of Automatic Speech Recognition for Military Applications," IEEE Trans ASSP, Vol. ASSP-25, pp. 310-322, (1977).

7. Bell, C. G., H. Fujisaki, J. M. Heinz, K. N. Stevens, and A. S. House, "Reduction of Speech Spectra by Analysis-by-Synthesis Techniques," J Acoust Soc Am, Vol. 33, pp. 1725-1736, (1961).

8. Bellamy, John, Digital Telephony, Ch. 3, John Wiley and Sons, New York, NY (1982).

9. Berry, Bob, E-Systems, Garland Div., Personal Communication.

10. Blankenship, P. E., MIT Lincoln Laboratory, Personal Communication.

11. Blankenship, P. E. and M. L. Malpass, "Frame-Fill Techniques for Reducing Vocoder Data Rates," Lincoln Laboratory Technical Report, No. 556, (1981).

12. Campanella, S. J. and G. S. Robinson, "A Comparison of Walsh and Fourier Transforms for Application to Speech," Proc 1971 Symp Walsh Functions, pp. 199-205, (1971).

13. Carr, Jim, E-Systems, Garland Div., Personal Communication.

14. Cheung, R. S. "Real-Time Implementation of a 9600 BPS Subband Coder with Time-Domain Harmonic Scaling," ICASSP '82, pp. 208-211, (1982).

15. Childers, D. G., D. P. Skinner, and R. C. Kemerait, "The Cepstrum: A Guide to Processing," Proc IEEE, Vol. 65, pp. 1428-1443, (1977).

16. Christensen, R. L., W. J. Strong, and E. P. Palmer, "A Comparison of Three Methods of Extracting Resonance Information from Predictor-Coefficeint Coded Speech," IEEE Trans ASSP, Vol ASSP-24, pp. 8-14, (1976).

17. Cohn, D. L. and J. L. Melsa, "The Residual Encoder--An Improved ADPCM System for Speech Digitization," IEEE Trans Comm, Vol. COM-23, pp. 935-941, (1975).

18. Crain, Vicki, Motorla, Inc., Personal Communication.

19. Crochiere, R. E. and J. L. Flanagan, "Current Perspectives in Digital Speech," IEEE 1982 Int'l Comm Conf, Vol. 3, pp. 3G.3.1-3G.3.5, (1982).

20. Dalton, John, ITT, Personal Communication.

21. Dudley, Homer, "Phoetic Pattern Recognition Vocoder for Narrow-Band Speech Transmission," J Acoust Soc Am, Vol. 30, pp. 733-739, (1958).

22. E-Systems, Garland Division. Installation, Operation, and Maintenance Instructions--CV-3333A/U Audio Digital Converter.

23. ________, Installation, Operation, and Maintenance Instructions--CV-3670/A LPC Speech Processor and C-10085/A Remote Control Unit.

24. ________, "CV-3333A/U Audio-Digital Converter," Specification Sheet.

25. ________, "CV-3333/U Digital Channel Speech Processor and CV-3670/A Linear Predictive Coding (LPC) Speech Processor," Fact Sheet.

26. ________, "CV-3333/U Mil Spec Digital Speech Processor," Specification Sheet.

27. ________, "CV-3670/A Airborne Digital Speech Processor," Specification Sheet.

28. ________, "Linear Predictive Coder Model LPC-24 Digital Speech Processor," Specification Sheet.

29. Feldman, J. A., E. M. Hofstetter, and M. L. Malpass. "A Compact, Flexible, LPC Vocoder Based on a Commercial Signal Processing Microcomputer," IEEE J Solid-State Circuits, Vol. SC-18, pp. 4-9, (1983).

30. Fette, Bruce, Motorola, Inc., Personal Communication.

31. Flanagan, J. L., "A Difference Limen for Vowel Formant Frequency," J Acoust Soc Am, Vol. 27, pp. 613-617, (1955).

32. ________, "Digital Voice Processing for Transmission and Storage," GLOBECOM '82, pp. E6.1.1-E6.1.3, (1982).

33. ________, Speech Analysis, Synthesis, and Perception, Academic Press, Inc., New York, NY, (1965).

34. Flanagan, J. L., M. R. Schroder, B. S. Atal, R. G. Crochiere, N. S. Jayant, and J. M. Tribolet, "Speech Coding," IEEE Trans Comm, Vol. COM-27, pp. 710-735, (1979).

35. Flanagan, J. L. and R. M. Golden, "Phase Vocoder," Bell Sys Tech J, Vol. 45, pps. 1493-1509, (1966).

36. Fransen, L. J., "2400- to 800-b/s LPC Rate Converter," NRL Report 8716, (1983).

37. Fransen, Larry, Naval Research Laboratory, Personal Communication.

38. Fussell, J. W., "A Differential Linear Predictive Voice Coder for 1200 bps," ICASSP '81, pp. 24-27, (1981).

39. Garner, J., T. McChesney, and M. Glancy, "Advanced Narrowband Digital Voice Terminal," reprint, Signal Magazine, (1982).

40. Gibson, J. D., J. L. Melsa, S. K. Jones," Digital Speech Analysis Using Sequential Estimation Techniques," IEEE Trans ASSP, Vol. ASSP-23, pp. 362-369, (1975).

41. Gold, Bernard, "A Computer Program for Pitch Extraction," J. Acoust. Soc. Am., Vol. 34, pp. 916-921, (1962).

42. ________, "Digital Speech Networks," Proc IEEE, Vol. 65, pp. 1636-1658, (1977).

43. ________, "Experiments with a Pattern-Matching Channel Vocoder," ICASSP '81, pp. 32-34, (1981).

44. ________, MIT Lincoln Laboratory, Personal Communication.

45. Gold, B. and C. M. Rader, "The Channel Vocoder," IEEE Trans AU, Vol AU-15, pp. 148-161, (1967).

46. ________, "Systems for Compressing the Bandwidth of Speech," IEEE Trans AU, Vol. AU-15, pp. 131-136, (1967).

47. Gold, B. and J. Tierney, "Vocoder Analysis Based on Properties of the Human Auditory System," Lincoln Laboratory Technical Report, No. 670, (1983).

48. Gold, B., P. G. Blankenship, and R. J. McAulay, "New Applications of Channel Voocders," IEEE Trans ASSP, Vol. ASSP-29, pp. 13-22, (1981).

49. Goldberg, A. J., "Analog to Digital Conversion of Speech--A Survey and Overview," Defense Communications Agency Report, No. TM 3-70, (1970).

50. Gray, A. H., Jr., and J. D. Markel, "Quantization and Bit Allocation in Speech Processing," IEEE Trans ASSP, Vol. ASSP-24, pp. 459-473, (1976).

51. GTE Systems, "Model DVT-2000G Telephone Terminal with User Furnished Crypto Specifications," Product Information Sheet.

52. ________, "Models MRD-2000G and UVD-2000 Voice Digitizers," Product Information Sheet.

53. ________, "Multiple Rate Voice Terminal" Technical Product Information Summary.

54. Gutman, L. L., Air Force Wright Aeronautical Laboratories, Avionics Laboratory, Personal Communication.

55. Heiler, Josef, "Optimized Frame Selection for Variable Frame Rate Synthesis," ICASSP '82, pp. 586-588, (1982).

56. Hess, W. J., "A Pitch-Synchronous Digital Feature Extraction System for Phonemic Recognition of Speech," IEEE Trans ASSP, Vol. ASSP-24, pp. 14-25, (1976).

57. Hofstetter, E. M., J. Tierney, and O. Wheeler, "Microprocessor Realization of a Linear Preditive Vocoder," IEEE Trans ASSP, Vol. ASSP-25, pp. 379-387, (1977).

58. Holmes, J. N., "The JSRU Channel Vocoder," IEEE Proc, Vol. 129, pp. 53-60, (1980).

59. Huggins, W. H., "A Note on Autocorrelatin Analysis of Speech Sounds," J Acoust Soc Am Vol. 26, pp. 790-792, (1954).

60. Itakura, F. and S. Saito, "Analysis Synthesis Telephony Based on the Maximum Likelihood Method," Proc Sixth Int'l Congr Acoust, pp. C-17 - C-20, (1968).

61. ITT, Defense Communications Div., "Advanced Narrowband Digital Voice Terminal ANDVT," Product Information Sheet.

62. Jayant, N. S., "Delta Modulation of Pitch, Formant, and Amplitude Signals for the Synthesis of Voiced Speech," IEEE Trans AU, Vol. AU-21, pp. 135-140, (1973).

63. Joint Tactical Communications Office, "Performance Specification for the ANDVT Tactical Terminal (TACTERM)," Specification No. TT-B1-4210-0087B, 30 Jan 1981.

64. ________, "Appendix 1--Performance Specification for LPC Processor," Specification No. TT-BA-4210-0087, 31 May 1978.

65. Kabriski, M. J., Air Force Institute of Technology, EE Dept., Personal Communication.

66. Kang, G. S., Naval Research Laboratory, Personal Communication.

67. Kang, G. S., and D. C. Coulter, "600-Bit-Per-Second Voice Digitizer (Linear Predictive Formant Vocoder)," NRL Interim Report, No. 8043, (1976).

68. Kizer, L. R., Air Force Institute of Technology, EE Dept., Personal Communication.

69. Kopec, G. G., A. V. Oppenheim, and J. M. Tribolet, "Speech Analysis by Homomorphic Prediction," IEEE Trans ASSP, Vol. ASSP-25, pp. 40-49, (1977).

70. Kulya, V. I., "Experimental Investigation of the Correlation Relations in the Speech Spectrum and a Comparison of Some Variants of Orthogonal Vocoder," Telecommunications, Vol. 18, pp. 39-50, (1964).

71. Lacroix, A. and B. Makai, "A Vocoder Scheme for Very Low Bit Rates (Quality Evaluation)," ICASSP '82, pp. 618-621, (1982).

72. Langston, J. L., Texas Instruments, Personal Communication.

73. Lin, K. S. and Y. L. Tsui, "LPC Compressed Speech at 850 Bits-Per-Second," ICASSP '81, pp. 5-7, (1981).

74. Makhoul, John, "Linear Prediction: A Tutorial Review," Proc IEEE, Vol. 63, pp. 561-580, (1975).

75. ________, "Spectral Analysis of Speech by Linear Prediction," IEEE Trans AU, Vol. AU-21, pp. 140-148, (1973).

76. ________, "Spectral Linear Prediction: Properties and Applications," IEEE Trans ASSP, Vol. ASSP-23, PP. 283-297, (1975).

77. Makhoul, J. and R. Viswanathan, "Quantization Properties of Transmission Parameters in Linear Predictive Systems," BBN Report, No. 2800, (1974).

78. Maksym, J. N., "Real-Time Pitch Extraction by Adaptive Prediction of the Speech Waveform," IEEE Trans AU, Vol. AU-21, pp. 149-154, (1973).

79. Manley, H. J., GTE Systems, Personal Communication.

80. Manley, H. J., and H. Schaffer, C. R. Howard, and J. O'Brien, "All-Digital Vocoder Techniques," Technical Report AFAL-TR-69-302, Part I, Sylvania Electronic Systems, (1969).

81. Manley, H. J., H. Shaffer, C. R. Howard, J. O'Brien, and R. Freudberg, "All Digital Vocoder Techniques," Technical Report AFAL-TR-67-302, Part II, Sylvania Electronic Systems, (1971).

82. Markel, J. D., "Applications of a Digital Inverse Filter for Automatic Formant and $F_0$ Analysis," IEEE Trans AU, Vol. AU-21, pp. 154-160, (1973)

83. ________, "Digital Inverse Filtering--A New Tool for Formant Trajectory Estimation," IEEE Trans AU, Vol. AU-20, pp. 129-139, (1972)

84. Markel, J. D., and A. H. Gray, Jr., "On Autocorrelation Equations as Applied to Speech Analysis," IEEE Trans AU, Vol. AU-21, pp. 69-79, (1973).

85. Mathews, M., M. Miller, and E. David, Jr., "Pitch Synchronous Analysis of Voiced Sounds," J Acoust Soc Am, Vol. 33, pp. 179-186, (1961).

86. Matsuyama, Y. and R. M. Gray, "Voice Coding and Tree Encoding Speech Compression Systems Based Upon Inverse Filter Matching," IEEE Trans Comm, Vol. COM-30, pp. 711-720, (1982)

87. McAuly, R. J., "A Low-Rate Vocoder Based on an Adaptive Subband Formant Analysis," ICASSP '81, pp. 28-33, (1981).

88. McCandless, S. P., "An Algorithm for Automatic Formant Extraction Using Linear Prediction Spectra," IEEE Trans ASSP, Vol. ASSP-22, pp. 135-141, (1974).

89. Motorola, Inc., Communications Division, Product Information Report.

90. Niederjohn, R. J., "A Mathematical Formulation and Comparison of Zero Crossing Analysis Techniques Which Have Been Applied to Automatic Speech Recognition," IEEE Trans ASSP, Vol. ASSP-23, pp. 373-380, (1975).

91. Noll, P., "A Comparitive Study of Various Quantization Schemes for Speech Encoding," Bell Sys Tech J, Vol. 54, pp. 1597-1614, (1975).

92. O'Donnell, J. J., "A System for Very Low Data Rate Speech Communication," ICASSP '81, pp. 8-11, (1981).

93. Olive, J. P., "Automatic Formant Tracking by a Newton-Raphson Technique," J Acoust Soc Am, Vol. 50, pp. 661-670, (1971).

94. Oppenheim, A. V., "Speech Analysis-Synthesis System Based on Homomorphic Filtering," J Acoust Soc Am, Vol. 45, pp. 458-465, (1969).

95. ________, Ed., Applications of Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, NJ, (1978).

96. Oppenheim, A. V. and R. W. Schafer, Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, NJ, (1975).

97. Papamichalis, P. G. and G. R. Doddington, "Time Encoding of LPC Roots," ICASSP '82, pp. 589-592, (1982).

98. Paul, D. B., "A 500-800 bps Adaptive Vector Quantization Vocoder Using a Perceptually Motivated Distance Measure," GLOBECOM '82, pp. E6.3.1-E6.3.4, (1982).

99. ________, "The Spectral Envelope Estimation Vocoder," IEEE Trans ASSP, Vol. ASSP-29, pp. 786-794.

100. ________, MIT Lincoln Laboratories, Private Communication.

101. Paul, D. B. and P. E. Blankenship, "Two Distance Measure-Based Vocoder Quantization Algorithms for Very-Low Data Rate Aplications: Frame-Fill and Spectral Vector Quantization," ICCR, pp. 3G.5.1-3G.5.6, (1982).

102. Portnoff, M. R. "Short-Time Fourier Analysis of Sampled Speech," IEEE Trans ASSP, Vol. ASSP-29, pp. 364-373, (1981).

103. Rabiner, L. R., "Digital-Formant Synthesizer for Speech Synthesis Studies," J Acoust Soc Am, Vol. 43, pp. 822-828 (1968).

104. Rabiner, L. R. and B. Gold, Theory and Applications of Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, NJ, Chapter 12, (1975).

105. Rebolledo, G., R. M. Gray, and J. P. Burg, "A Multirate Voice Digitizer Based Upon Vector Quantization," IEEE Trans Comm, Vol. COM-30, pp. 721-727, (1982).

106. Reference Data for Radio Engineers, 6th Ed., Howard W. Sams and Co., Inc., Indianapolis, IN, (1975).

107. Roucos, S., J. Markhoul, and R. Schwartz, "A Variable-Order Markov Chain for Coding of Speech Spectra," ICASSP '82, pp. 582-585, (1982).

108. Roucos, S., R. Schwartz, and J. Makhoul, "Vector Quantization for Very-Low-Rate Coding of Speech," GLOBECOM '82, pp. 1074-1078, (1982).

109. Sabin, M. J. and R. M. Gray, "Product Code Vector Quantizers for Speech Waveform Coding," GLOBECOM '82, pp. E6.5.1 - E6.5.5, (1982).

110. Sambur, M. R., "An Efficient Linear-Prediction Vocoder," Bell Sys Tech J, Vol. 54, pp. 1693-1723, (1975).

111. Schafer, R. W. and L. R. Rabiner, "Design and Simulation of a Speech Analysis-Synthesis System Based on Short-Time Fourier Analysis," IEEE Trans AU, Vol. AU-21, pp. 165-174, (1973).

112. ________, "System for Automatic Formant Analysis of Voiced Speech," J Acoust Soc Am, Vol. 47, pp. 634-648, (1970).

113. Schmidt-Nielson, A. and S. G. Everett, "A Conversation Test for Comparing Voice Systmes Using Working, Two-Way Communications Links," IEEE Trans ASSP, Vol. ASSP-30, pp. 853-863, (1982).

114. Schroeder, M. R., "Correlation Techniques for Speech Bandwidth Reduction," J Audio Eng Soc, Vol. 10, pp. 163-166, (1962).

115. ________, "Vocoders: Analysis and Synthesis of Speech," Proc IEEE, Vol. 54, pp. 720-734, (1966).

116. Schroeder, M. R., J. L. Flanagan, and E. A. Lundry, "Bandwidth Compression of Speech by Analytic-Signal Rooting," Proc IEEE, Vol. 55, pp. 396-401, (1967).

117. Schwartz, M. Information Transmission, Modulation, and Noise, 3rd Ed., McGraw-Hill Book Co., New York, NY, (1980).

118. Seelandt, K. G., "Computer Analysis and Recognition of Phoneme Sounds in Connected Speech," MS Thesis GE/EE 81D-53, AFIT, (1981).

119. Segota, Anton, Rome Aeronautical Development Center (RADC), Hanscom AFB, Personal Communication.

120. Shum, F. Y. Y., A. R. Elliott, and W. O. Brown, "Speech Processing with Walsh-Hadamard Transforms," IEEE Trans AU, Vol. AU-21, pp. 174-179, (1973).

121. Singer, E., "A Comparitive Study of Narrowband Vocoder Algorithms in Air Force Operational Environments Using the Diagnostic Rhyme Test," LL Technical Report, No. 590, (1982).

122. ________, "The Effects of Microphones and Facemasks on LPC Vocoder Performance," LL Technical Report, No. 584, (1981).

123. Smith, C. P., RADC, Hanscom AFB, Personal Communication.

124. "State of the Art Speech Compression and Digitization," ESD Technical Documentary Report, ESD-TPR-64-460, (1964).

125. Steigerwald, Silvi, RADC, Griffiss AFB, Personal Communication.

126. Steiglitz, K., G. Winham, and J. Petzinger, "Pitch Extraction by Trigonometric Curve Fitting," IEEE Trans ASSP, Vol. ASSP-23, pp. 321-323, (1975).

127. Stewart, L., R. M. Gray, and Y. Linde, "The Design of Trellis Waveform Coders," IEEE Trans Comm, Vol. COM-30, pp. 702-710, (1982).

128. Strong, W. J., "Machine-Aided Formant Determination for Speech Synthesis," J Acoust Soc Am, Vol. 41, pp. 1434-1442, (1967).

129. Sudhoff, R. I., NAVLEX, Personal Communication.

130. Transmission Systems for Communications, 3rd ed., Western Electric Co., Inc., Techncial Publications, Winston-Salem, NC, (1964).

131. "Test Report for Development Test and Evaluation of the Advanced Narrowband Digital Voice Terminal (ANDVT)," extract, TRI-TAC JTO, (1983).

132. Un, C. K., "A Low-Rate Digital Formant Vocoder," IEEE Trans Comm., Vol. COM-26, pp. 344-355, (1978).

133. Un, C. K. and D. T. Magill, "The Residual-Excited Linear Prediction Vocoder with Transmission Rate Below 9.6 kbits/s," IEEE Trans Comm, Vol. COM-23, pp. 1466-1474, (1975).

134. Un, C. K. and J. R. Lee, "On Spectral Flattening Techniques in Residual-Excited Linear Prediction Vocoding," ICASSP '82, pp. 216-219, (1982).

135. Vetrano, Parker, RADC, Hanscom AFB, Personal Communication.

136. Viswanathan, V. R., J. Makhoul, "Quantization Properties of Transmission Parameters in Linear Preditive Systems," IEEE Trans ASSP, Vol. ASSP-23, pp. 309-321, (1975).

137. Viswanathan, V. R., J. Makhaul, R. M. Schwartz, and A. W. F. Huggins, "Variable Frame Rate Transmission: A Review of Methodology and Application Narrow-Band LPC Speech Coding," IEEE Trans Comm, Vol. COM-30, pp. 674-686, (1982).

138. Viswanathan, V. R., M. Berouti, A. Higgins, W. Russell, "A Harmonic Deviations Linear Prediction Vocoder for Improved Narrowband Speech Transmission," ICASSP '82, pp. 610-613, (1982).

139. Voiers, W. D., "Antecedents of Speech Acceptability," International Symposium on Measurements in Telecommunications, reprint, (1977).

140. ________, "Diagnostic Acceptability Measure for Speech Communication Systems," ICASSP '77, reprint, (1977).

141. ________, "Diagnostic Evaluation of Speech Coding Systems," Defense Communications Agency Technical Report, DCA 100-77-C-0003, (1978).

142. ________, "Diagnostic Evaluation of Speech Intelligibility," Benchmark Papers in Acoustics, Vol. 11, [(M. Hawley, ed.), Dowden, Hutchison, and Ross, Straudsburg, PA, (1977)].

143. ________, "Evaluating Processed Speech Using the Diagnostic Rhyme Test," Speech Technology, reprint, (1983).

144. ________, "Some Thoughts on the Standardization of Psychological Measures of Speech Intelligibility and Quality," Proceedings of the Workshop on Standardization for Speech I/O Technology, reprint, (1982).

145. ________, "Uses, Limitations, and Interrelations of Present-Day Intelligibility Tests," National Electronics Conference Proceedings, Vol. 35, pp. 251-255, (1981).

146. ________, Dynastat, Inc., Personal Communication.

147. Weldon, Wayne, E-Systems, Inc., Personal Communication.

148. Wilkins, Frank, GTE Systems, Personal Communcation.

149. Wong, D. Y. "Vector/Matrix Quantization for Narrow-Bandwidth Digital Speech Compression," RADC Final Report, RADC-TR-82-246, (1982).

150. Wong, D. Y. and B. H., Juang, "Voiced Coding at 800 BPS and Lower Data Rates with LPC Vector Quantization," ICASSP '82, pp. 606-609, (1982).

151. Wong, D. Y., B. H. Juang, and A. H. Gray, Jr., "An 800 bit/s Vector Quantization LPC Vocoder," IEEE Trans ASSP, Vol. ASSP-30, pp. 770-780, (1982).

152. ________, "Recent Developments in Vector Quantization for Speech Processing," ICASSP '81, pp. 1-4, (1981).

## APPENDIX A1

### SPEECH BANDWIDTH COMPRESSION GAINS IN LPI COMMUNICATIONS

An LPI communications system is an attempt to maximize the likelihood of correct reception of voice and/or data transmissions by an intended receiver while minimizing the likelihood that an intercept receiver will be able to detect the communication process in progress. This is accomplished through the use of a communications system incorporating any or all of the techniques mentioned in Chapter I. The goal is to operate at absolutely the lowest possible RF energy level necessary to convey the information to the intended receiver. As indicated by the list of technologies under investigation, this process will probably be an adaptive one, continuously changing the operating characteristics within a closed loop communications situation. The vocoder used for speech bandwidth compression is expected to yield significant gains towards the LPI capabilities of the composite LPI system.

A vocoder fits into the communications link as previously described in Chapter II. The vocoder will be utilized to replace the PCM, ADPCM, DM, etc., modems now used in digital radio systems in the source encoding portion of the system (see Figure 2-1). The source encoder outputs the vocoder data at a specific rate, $R_v$. The encryption process rearranges the data or combines it with a known (to the

receiver) sequence of information to prevent unwanted interpretation or decoding of the data. The channel encoder generally adds bits to the data stream proportional to the incoming rate for error protection/correction purposes. The actual number of bits added depends upon the amount of error protection desired. At this point, the output is the bit rate, R, which is the system data rate. Therefore, the vocoder establishes the overall system bit rate.

The communications channel is fixed such that a value termed processing gain, PG, is the variable directly affected by the data rate. This in turn affects the maximum coherent reception and interception ranges. By starting with the range equations, the processing gain can be derived to show the LPI gains achieved by reducing the data rate.

The coherent reception range, $R_R$, is defined as the maximum range at which an intended receiver may detect the communication signal. The interception range, $R_I$, is defined as the maximum range at which an unintended receiver may detect the communication emissions. The free space equations expressing those ranges in system parameters are given below (54).* The range for intended reception is:

$$R_R^2 = \frac{P_T t_0 G_{TR} G_{RR} \lambda^2}{(4\pi)^2 k T_R F_R L_R D_R} , \qquad \text{(A1-1)}$$

---

*kTF in the denominators is actually $kT_s[\frac{Tc}{T_s} + (F-1)]$ but assuming ideal conditions $T_c = T_s$.

and the range for interception is:

$$R_I^2 = \frac{P_T G_{TI} G_{RI} \lambda^2}{(4\pi)^2 k T_I F_I L_I d_I B_I} ; \tag{A1-2}$$

where $P_T$ is the transmitter peak power, $t_0$ is the coherent integration time, $G_T$ and $G_R$ are the appropriate transmit and receive antenna gains, respectively, $\lambda$ is the carrier frequency wave length, k is Boltzman's constant, T is the appropriate receiver noise temperature, F is the appropriate receiver noise figure, L is the appropriate receiver loss figure, d is the appropriate signal-to-noise (SNR) power ratio required for detection (after coherent processing), and B is the appropriate receiver noise bandwidth.

Multiplying the numerator and demoninator of (A1-1) by the receiver bandwidth $B_R$, the time-bandwidth product, $D_R$, where:

$$D_R = t_0 B_R \tag{A1-3}$$

is obtained. Equation (A1-1) can now be rewritten as:

$$R_R^2 = \left[\frac{G_{TR} G_{RR} \lambda^2}{(4\pi)^2 k T_R F_R L_R d_R B_R}\right] D_R P_T . \tag{A1-4}$$

Letting the terms in brackets equal a constant, $M_R$, (A1-4) can be reduced to:

$$R_R^2 = M_R D_R P_T . \tag{A1-5}$$

Equation (A1-2) can be rearranged as:

$$R_I^2 = \left[\frac{G_{TI} G_{RI} \lambda^2}{(4\pi)^2 k T_I F_I L_I d_I B_I}\right] P_T \tag{A1-6}$$

or

$$R_I^2 = M_I P_T \ . \tag{A1-7}$$

The maximum operating range, $R_R$, is determined by system and mission requirements. For this research effort, worst-case conditions are assumed giving the situation where the coherent receiver and the interception receiver have identical system characteristics. This means that $M_R = M_I$. *In actuality*, these values will not be equal but because the terms of the expression are essentially constants, they will be fixed and, therefore, proportional. Equation simplification is the result of this assumption. Now, relating the coherent receiver range to the intercept receiver range yields:

$$\frac{R_R^2}{R_I^2} = \frac{M_R D_R P_T}{M_I P_T} \ , \tag{A1-8}$$

resulting in:

$$\frac{R_R^2}{R_I^2} = D_R \ . \tag{A1-9}$$

LPI communication techniques attempt to maximize this ratio. This ratio is generally much greater than one because the coherent receiver has some a priori knowledge about the signal being transmitted. It knows the format of the signal and how to process it to get maximum value. This leads to the concept of processing gain, PG, which is an alternative designation for the time-bandwidth product, $D_R$. Therefore,

$$D_R = PG \tag{A1-10}$$

is a function of the RF bandwidth and integration time or bit duration. Maximizing PG maximizes (A1-9) and, therefore, (A1-8). Relating (A1-10) to (A1-5), if $R_R^2$ is fixed and $M_R$ is a constant, then maximizing $D_R$ minimizes $P_T$, the required transmitter power. Reducing $P_T$ in (A1-7) reduces the interception range, $R_I^2$, making the system less susceptible to interception as desired.

The processing gains can often be more clearly seen in SNR equations. It can be shown that increasing PG decrease the RF SNR required. The processing gain or time-bandwidth product is:

$$PG = t_0 B \tag{A1-11}$$

or

$$PG = \frac{B}{R}\,, \tag{A1-12}$$

where B is the receiver RF bandwidth. The RF SNR is:

$$(\frac{S}{N})_{RF} = 10 \log (\frac{s}{n}) , \quad \text{(A1-13)}$$

where S is the RF signal power in dB, N is the RF noise power in dB, s is the RF signal power in watts, and n is the RF noise power in watts. Now:

$$s = e_b R \quad \text{(A1-14)}$$

and

$$n = n_0 B \quad \text{(A1-15)}$$

where $e_b$ is the energy per bit, R is the data rate, and $n_0$ is the noise per cycle of bandwidth, then:

$$(\frac{S}{N})_{RF} = 10 \log (\frac{e_b R}{n_0 B_{RF}}) . \quad \text{(A1-16)}$$

This can be rewritten as:

$$(\frac{S}{N})_{RF} = 10 \log [(\frac{e_b R}{n_0 B_{RF}})(\frac{R}{B_{bb}})] \quad \text{(A1-16)}$$

where $B_{bb}$ is the baseband bandwidth which is dependent upon the modulation scheme used. (Here BPSK is assumed so that $|R| = |B_{bb}|$ according to Nyquist's theory as given in [118]). Now, with

$$e_b R = s_{bb} \quad \text{(A1-17)}$$

and

$$n_0 B_{bb} = n_{bb} \, , \qquad \text{(A1-19)}$$

where $s_{bb}$ is the baseband signal power in watts and $n_{bb}$ is the baseband noise power in watts, the RF SNR is given by:

$$\left(\frac{S}{N}\right)_{RF} = 10 \log \left(\frac{s_{bb}}{n_{bb}} \frac{R}{B_{RF}}\right) \, . \qquad \text{(A1-20)}$$

This can be rewritten as:

$$\left(\frac{S}{N}\right)_{RF} = 10 \log \left(\frac{s_{bb}}{n_{bb}}\right) + 10 \log \left(\frac{R}{B_{RF}}\right) \qquad \text{(A1-21)}$$

or alternatively

$$\left(\frac{S}{N}\right)_{RF} = 10 \log \left(\frac{s_{bb}}{n_{bb}}\right) - 10 \log \left(\frac{B_{RF}}{R}\right) \, . \qquad \text{(A1-22)}$$

Since the data rate, R, is the reciprocal of the time, $t_0$, the second logarithmic term in (A1-22) is the time-bandwidth product in dB. The terms in (A1-22) can be expressed as:

$$10 \log \left(\frac{s_{bb}}{n_{bb}}\right) = \left(\frac{S}{N}\right)_{bb} \qquad \text{(A1-23)}$$

which is the baseband SNR in dB, and

$$10 \log \left(\frac{B_{RF}}{R}\right) = PG \qquad \text{(A1-24)}$$

which is the processing gain in dB. Now (A1-22) can be written as:

$$\left(\frac{S}{N}\right)_{RF} = \left(\frac{S}{N}\right)_{bb} - PG \ . \qquad \text{(A1-25)}$$

In (A1-25) the baseband SNR is determined by the information extraction circuits of a receiver. This value is the minimum SNR required to obtain a maximum probability of correctly interpreting the data. In the final analysis, (A1-25) shows that decreasing the bit rate increases PG and correspondingly a decrease in the SNR at the receiver front end is available to allow the transmitter output power to be reduced. Therefore, vocoder data rate reductions for speech bandwidth compression are directly applicable to LPI communication systems.

## APPENDIX A2

## DESCRIPTION OF THE SPEECH WAVE

> Speech is the acoustic end product of voluntary, formalized motions of the respiratory and masticatory apparatus. It is a motor behavior which must be learned. It is developed, controlled and maintained by the acoustic feedback of the hearing mechanism and by the kinesthetic feedback of the speech musculature. Information from these senses is organized and coordinated by the centeral nervous system and used to direct the speech function. impairment of either control mechanism usually degrades the performance of the vocal apparatus (33).

The purpose of speech analysis-synthesis systems is to efficiently encode the sounds of speech, transmit and receive this encoded signal, and decode the signal into perceptually significant sound. In order to best understand the various analysis--synthesis techniques, a reasonable understanding of the characteristics of the acoustic (speech) waveform is needed. The speech waveform can be characterized as the response of a slowly time varying system to either a quasi-periodic or a noise-like excitation.

> More specifically, the speech-production mechanism consists essentially of an acoustic tube, the vocal tract, excited by an appropriate source to generate the desired sound. In the case of <u>voiced</u> speech sounds, the excitation corresponds to a quasi-periodic pulse train representing the air flow through the cords as they vibrate. The fricative sounds are generated by forcing air through a constriction in the vocal tract, thereby creating turbulence, which produces a source of noise to excite the vocal tract (94).

The fricative sounds mentioned above are classed as unvoiced speech. This category also includes plosives. The voiced sounds

include vowels, nasals, and glides. The unvoiced sounds are the modulation of a noise-like excitation with the spectral envelope. The voiced sound (vowels) are the modulation of a more-or-less periodic excitation (vocal cord vibration with fundamental frequency equal to 1/excitation period) with the spectral envelope. Figures A2-1, A2-2, and A2-3 show a relatively long segment of speech showing samples of the various types of sounds. Figures A2-4 and A2-5 depict representations of these types of sounds.

The vocal tract can be assumed to be a linear time-varying system. Now, if the vocal-tract shape is fixed, or nearly so (slowly varying), the output of the system, the speech waveform, s(t), is approximated fairly accurately as the convolution of the given excitation source, c(t), and the vocal-tract impulse response, v(t), given as

$$s(t) = e(t) * v(t) \; . \qquad \text{(A2-1)}$$

In other words, the Fourier transform (spectrum) of the output is the product of the spectrums of the excitation function and the vocal-tract impulse response;

$$S(f) = E(f)V(f) \; . \qquad \text{(A2-2)}$$

The model is limited, and various difficulties can be noted. For example, the binary voicing decision does not provide for voiced fricatives (phonemes with simultaneous periodic and aperiodic excitation and a different filter for each excitation). The period of the periodic excitation may change rapidly or may only be quasi-periodic--either of which may cause sections of a short-term spectrum to be aperiodic. The filter itself may also change quite rapidly (99).

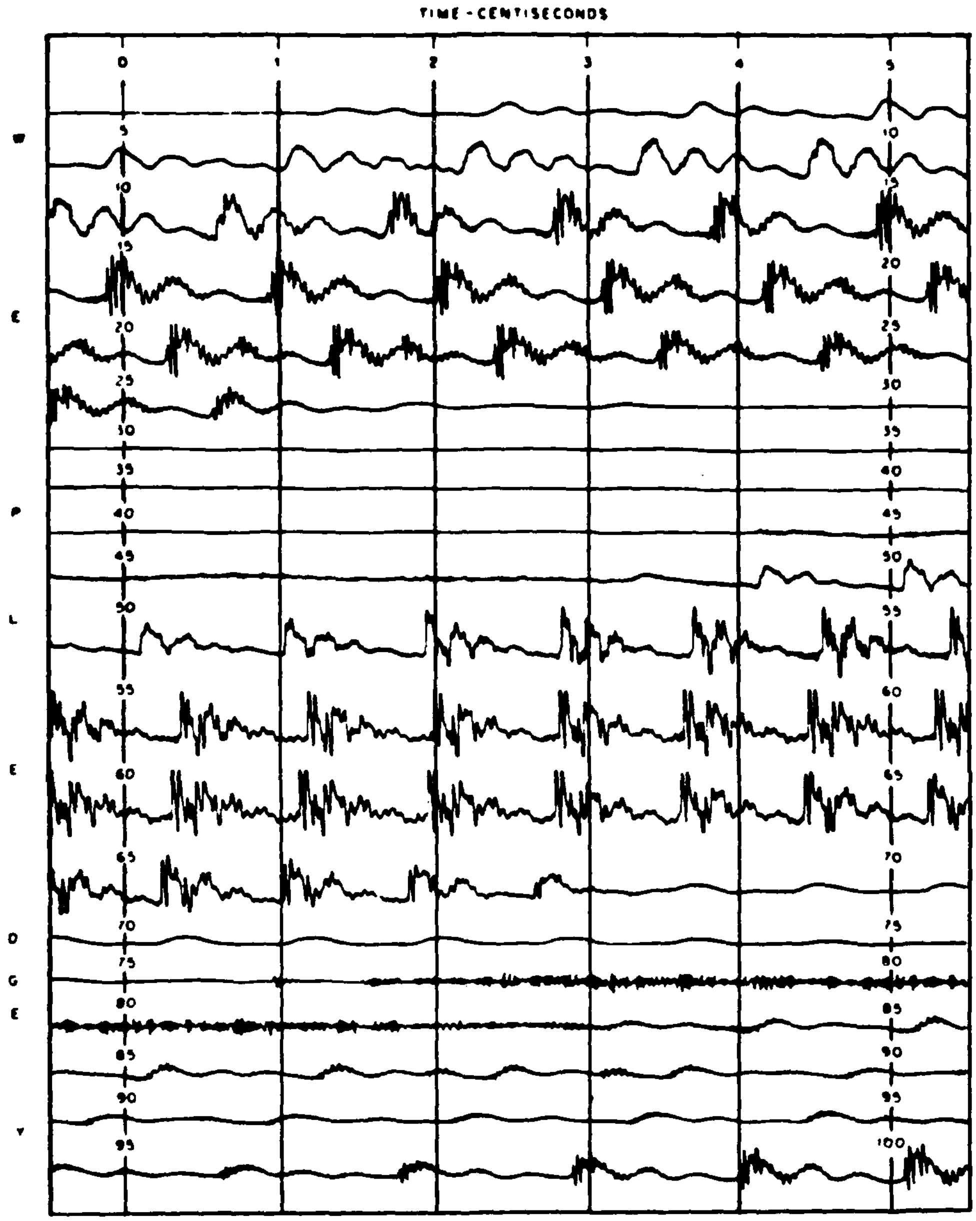



Figure A2-1. Beginning of speech waveform of the utterance "We pledge you some heavy treasure".
(Taken from reference 42, p. 1638)

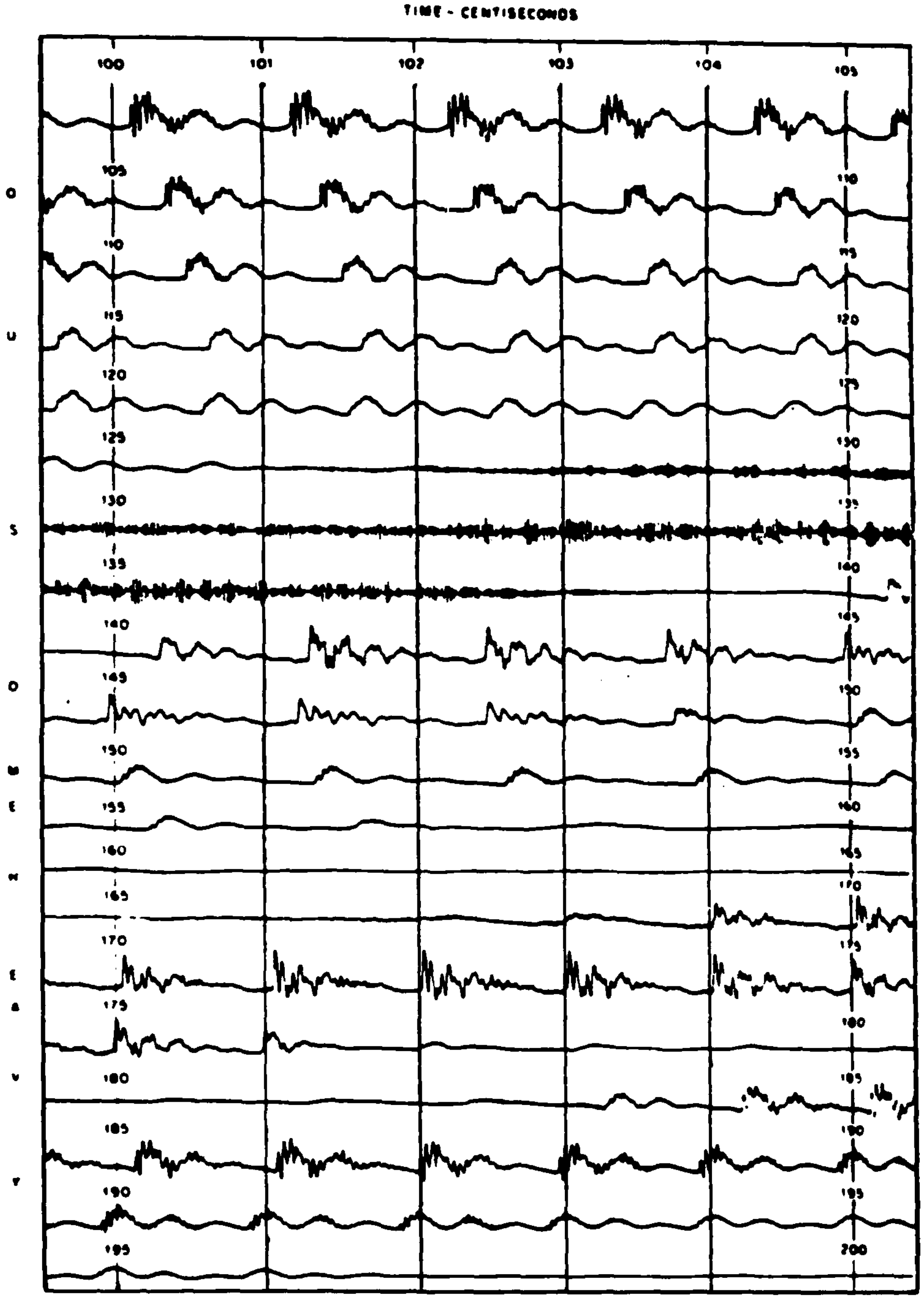



Figure A2-2. Continuation of speech waveform.
(Taken from reference 42, page 1638)

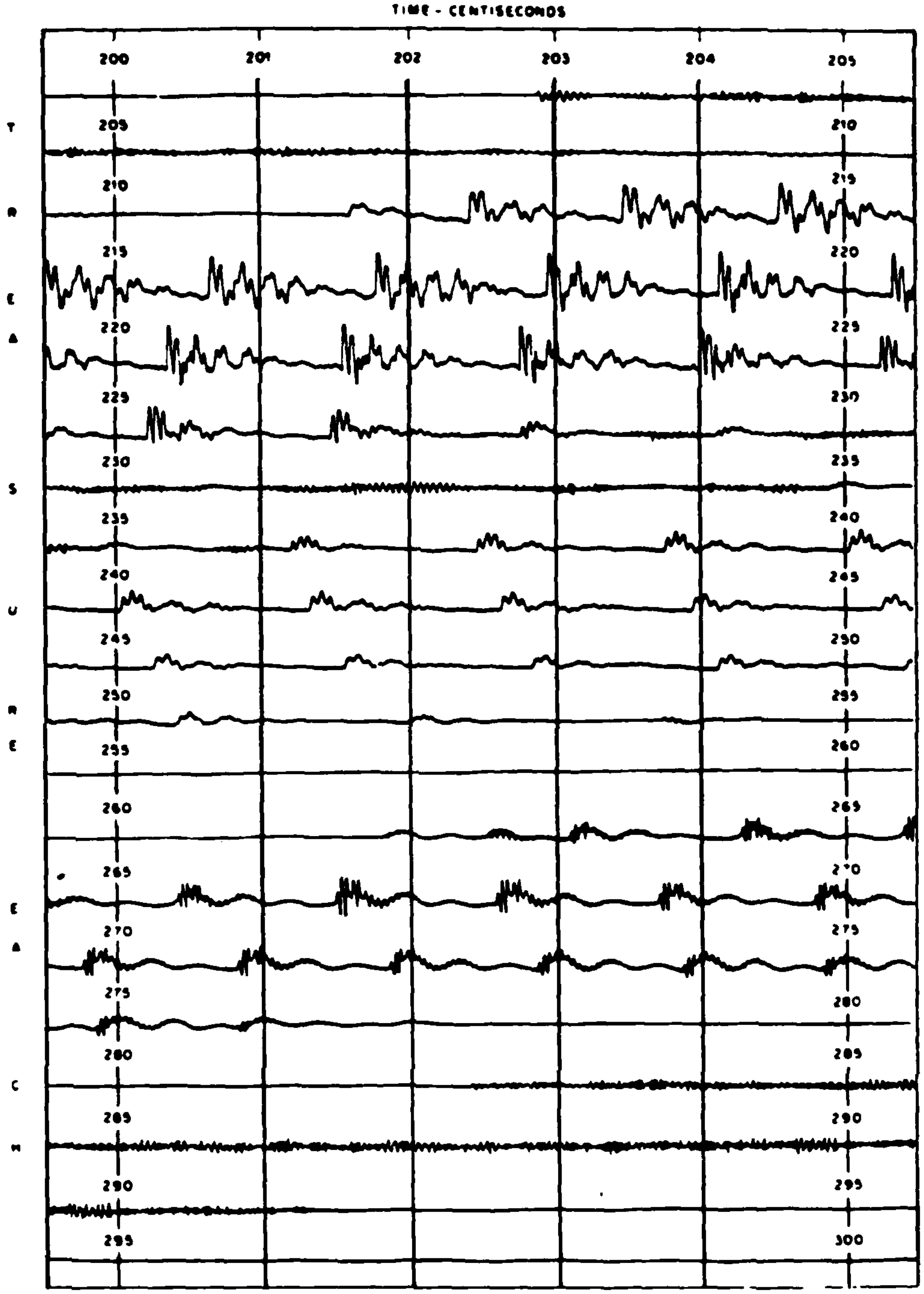



Figure A2-3. Continuation of speech waveform.
(Taken from reference 42, page 1639)

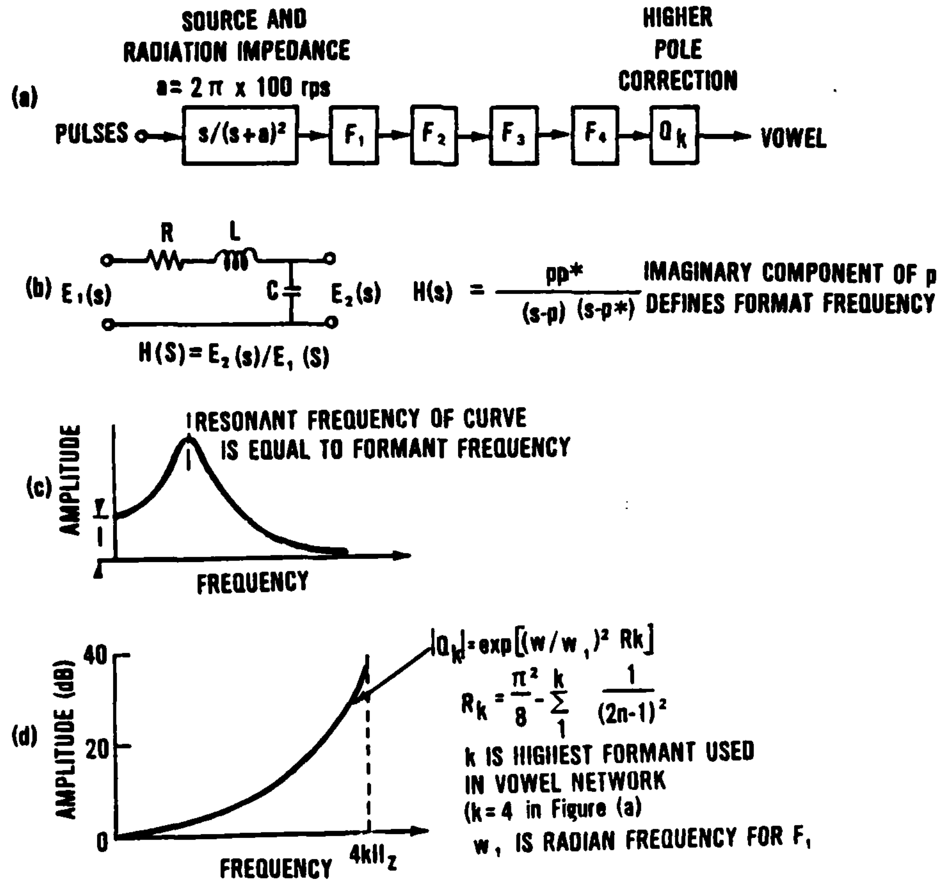



Figure A2-4. Details of network representation of vowels. (Taken from ref. 46, page 133)

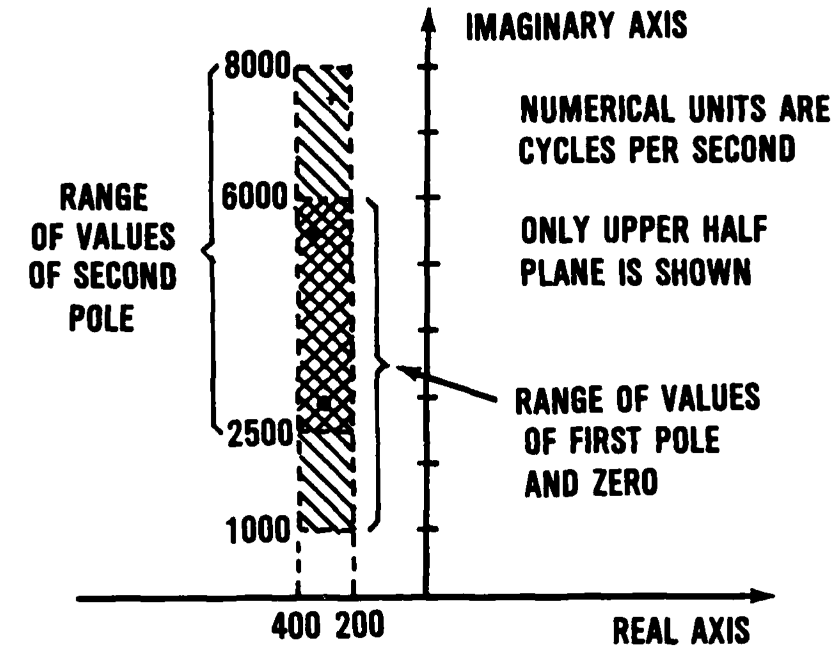



Figure A2-5. S-plane representation of voiceless plosives and voiceless fricatives.
(Taken from ref. 46, p. 133)

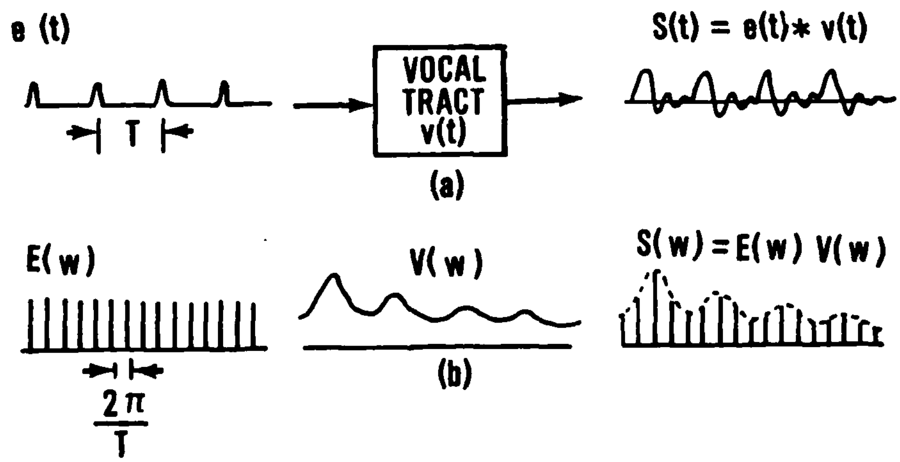



Figure A2-6. Model of speech production as the response of a quasi-stationary linear system; (a) time-domain characterization and (b) frequency-domain characterization.
(Taken from ref. 33, p. 121)

Figure A2-6 depicts the convolution and the spectrum product processes.

The spectrum of the vocal tract is a smooth, slow-varying function of frequency. The relative maximums, shown in Figure A2-6b, correspond to the resonant frequencies of the acoustic cavity, commonly called formant frequencies or just formants. The slowly-varying function is generally known as the speech envelope-structure or spectral envelope, G(f,t). The quasi-periodic excitation function has a period of approximately T. This produces a spectrum of pulses spaced $2\pi/T$ apart. The frequency $2\pi/T$ is the fundamental frequency or voice "pitch." The pitch is essentially constant (generally small variations) for an individual speaker but varies significantly between speakers. Pitch varies from approximately 50 Hz in adult men to about 400 Hz in women and children. The quasi-periodic function is referred to as the speech fine-structure, F(f,t).

All of the different characteristics mentioned above must be determined in some form or another. The pitch is determined separately from the voiced or unvoiced information, which is determined separately from the frequency content and signal amplitude.

## APPENDIX A3

## DESCRIPTION OF VOCODER TECHNIQUES

There are two different concepts in speech coding. These are waveform encoding and source encoding. Waveform encoding is essentially direct sampling and encoding of the speech waveform itself. This form attempts to completely model and quantize the wave and generally requires much higher data rates than source encoding. It is usually just an A/D conversion with appropriate resolution followed by the appropriate modulation scheme. Source encoding attempts to model some aspect of the vocalization/perception process, usually the vocal tract response function and the excitation function, at fairly low data rates. Vocoders are a form of source encoder. Figure A3-1 depicts the differences in the two forms. As can be seen in the figure, each form has its own advantages and applications. The most significant of these differences is the speech quality out of the receiver. Since the data rate for waveform encoders obviously exceeds the requirements of this research effort, they will not be discussed in this text. Source encoders are generally classified by the speech analysis techniques used. They are also often classified by their physical structure or sometimes by the parameters transmitted.

A source coder or vocoder attempts to analyze and characterize a speech waveform. The system must determine the type of sound,

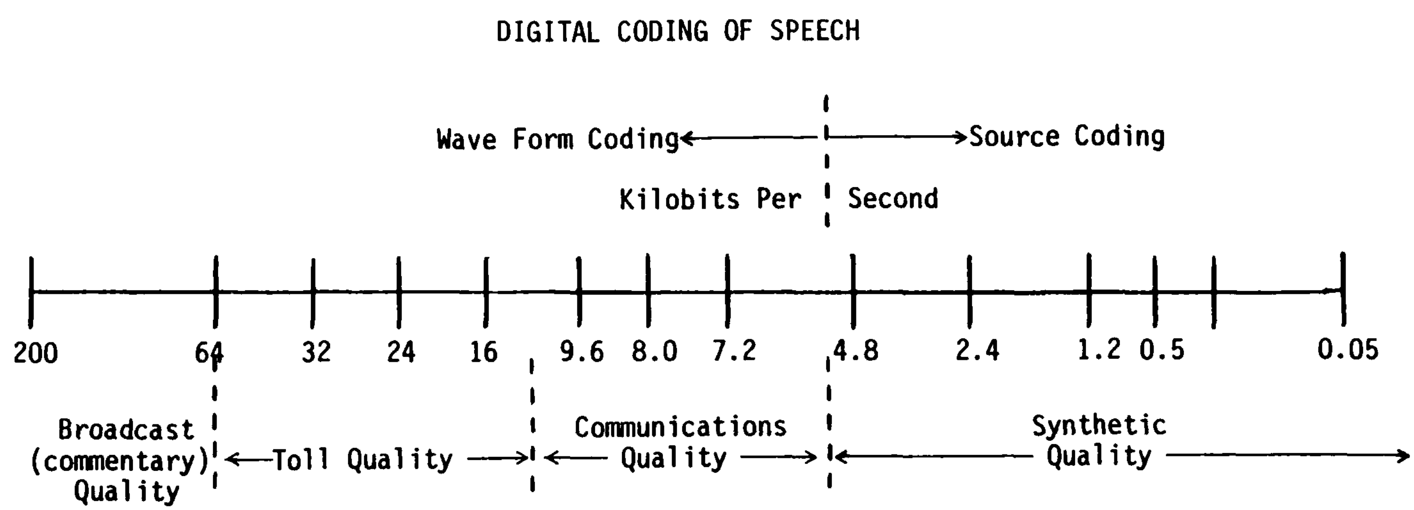



Figure A3-1. Spectrum of speech coding transmission rates (nonlinear scale) and associated quality. (Taken from ref. 34, p. 712)

voiced or unvoiced, the excitation period or pitch if voiced, and the frequency energy content of the signal. The vocoder must then characterize these parameters in such a manner that a synthesizer can input them and regenerate the speech wave as nearly identical as possible to the original wave analyzed. This appendix presents a detailed description of the seven major vocoder methods. These methods are the channel, formant, homomorphic, pattern-matching, phase, linear predictive coding, and spectral envelope estimation vocoders. Other minor techniques exist which are generally slight modifications or combinations of one or more of these major methods.

Channel Vocoder

The earliest vocoder dates back to 1928 when Homer Dudley of Bell Telephone Laboratories (115) sketched a device later to become known as the "vocoder." This early voice coder is the forerunner to what is now known as the spectrum channel vocoder or channel vocoder.

The channel vocoder is depicted in Figure A3-2. It consists of a number of channels. Each of the spectral channels shown here consists of a bandpass filter, a rectifer, and a low pass filter. The bandpass filters are established to continuously cover the desired speech bandwidth with a cut-off frequency usually between 3 kHz and 4 kHz. The end result of this series of channels is an estimate of the spectral envelope, $|G(f,t)|$. Because speech is a time-varying (only quasi-periodic) function (see Figures A2-1, A2-2, A2-3, and A2-6), infinite spectral analysis is not possible and is replaced by short-time spectral analysis. This method utilizes a time-window of

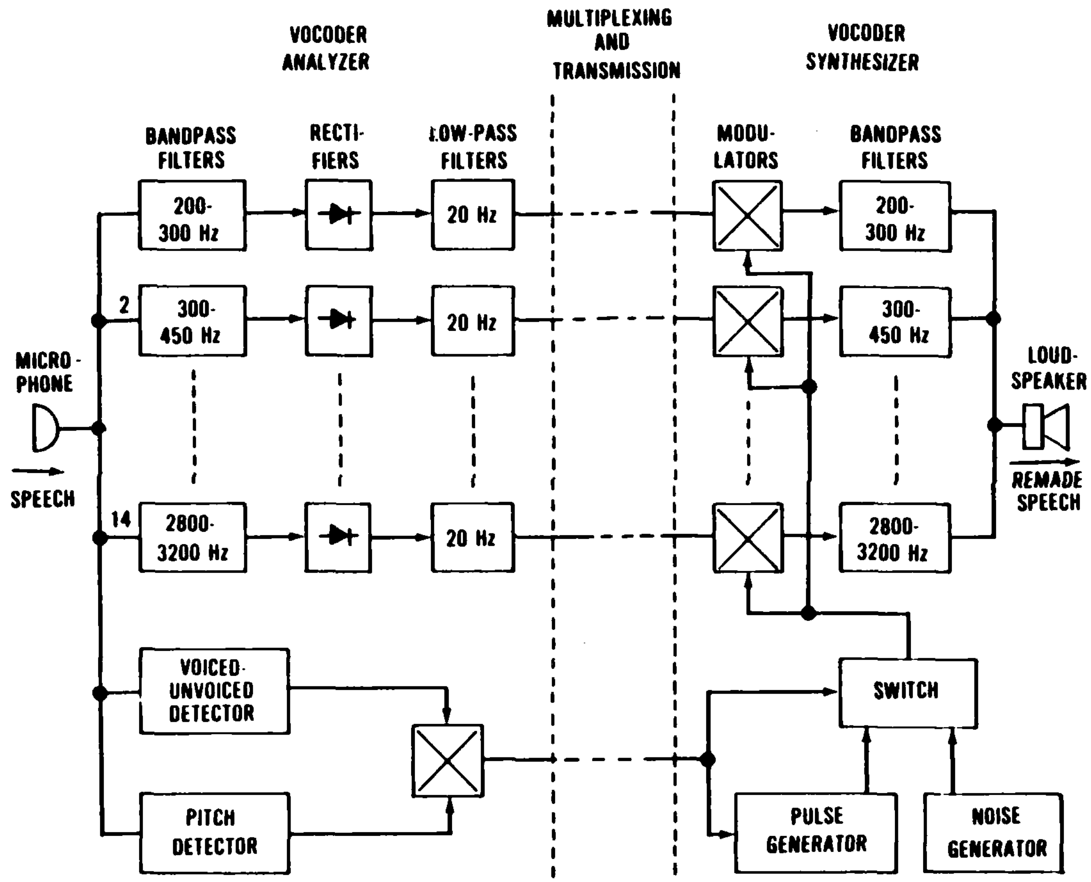



Figure A3-2. Block diagram of spectrum channel vocoder.
(Taken from ref. 115, p. 724)

duration approximately equal to the shortest speech sounds. This has been determined to be approximately 40 ms. Within the window, speech is assumed to be stationary. Equation A2-2 is written as:

$$S(f,t) = E(f,t)V(f,t) \; . \qquad \text{(A3-1)}$$

The spectrum can be expressed in terms of the spectral envelope, F(f,t), and the spectral fine-structure, G(f,t), as:

$$S(f,t) = F(f,t)G(f,t) \qquad \text{(A3-2)}$$

The channel vocoder performs a short-time Fourier analysis. The Fourier transform of a discrete signal is given by:

$$X(e^{jwT}) = \sum_{n=-\infty}^{\infty} x(nT)e^{-jwnT} \; . \qquad \text{(A3-3)}$$

The short-time transform is given as:

$$X(w,nT) = \sum_{r=-\infty}^{\infty} x(rT)h(nT-rT)e^{-jwrT} \; . \qquad \text{(A3-4)}$$

This equation is the infinite-time Fourier transform of the speech signal seen at time nT through a time window with response h(nT) as shown in Figure A3-3. The window response h(nT) transformed is $H(e^{jwT})$. This response is usually chosen to approximate the ideal low pass filter. Filters perform the analysis on the analog signal. If the bandpass filter response is limited to:

$$h_k(nT) = h(nT) \cos (w_k nT), \qquad \text{(A3-5)}$$

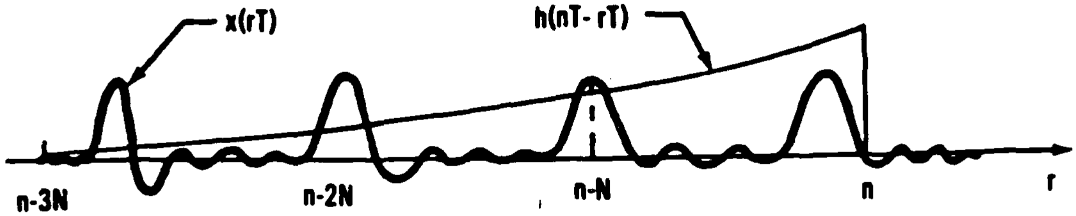



Figure A3-3. Representation of short-term spectrum analysis.
(Taken from ref. 104, p. 664)

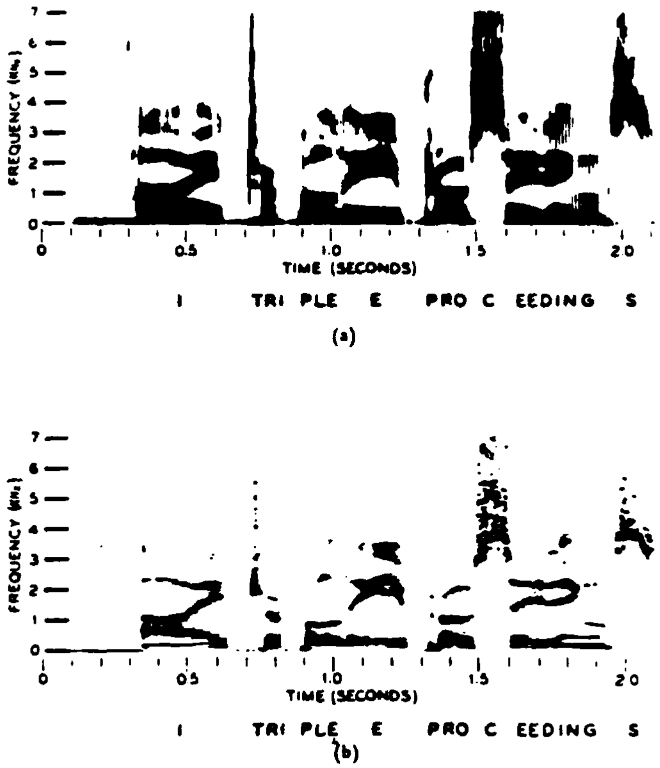



Figure A3-4. Speech spectrogram of "I triple E Proceedings", (a) wide-band analysis and (b) narrow-band analysis.
(Taken from ref. 115, p. 722)

then the bandpass output is given by:

$$y_k(nT) = \sum_{r=-\infty}^{n} x(rT)h(nT-rT) \cos [w_k(nT-rT)],$$

or

$$y_k(nT) = \mathrm{Re}[e^{jw_k nT} X(w_k, nT)]. \tag{A3-6}$$

The remaining channel makes the voiced/unvoiced (V/UV) decision and if voiced extracts the pitch. If the speech is voiced, the pitch signal sets the frequency of the pulse source and the V/UV signal selects the pulse source excitation to be modulated with the estimated envelope. If the sound is unvoiced, the "white" noise source of excitation is chosen.

Obviously, controlling the number of channels helps control the bit rate. It seems that 14 to 20 channels provides the optimum, practical number of vocoder channels. This generally provides bit rates between 1,000 bps and 4,800 bps with good quality (objectively speaking) speech.

## Formant Vocoder

The formant vocoder provides a somewhat more sophisticated approach to the spectral analysis of speech than the channel vocoder. The spectral envelope contains several prominent peaks around which the frequency components of the speech signal are grouped. These peaks are local resonances or resonant frequencies of the focal tract and are known as formants. Below 3 kHz there are usually three formants and below 4 kHz there are usually four or five formants. The statement "there are usually . . . below" indicates a general location

of the formants. These spectral peaks tend to shift with the production of different sounds. Figure A3-4 shows a sample of two speech spectrograms. The formants are frequency groupings appearing in the center of the more-or-less horizontal shadings. The narrow band analysis of (b) best indicates these formant frequencies. Figure A3-5 shows the formant and pitch structure of a small portion (the segment between 1.6 and 1.7 seconds) of Figure A3-4 in more detail.

The analyzer portion of the vocoder determines and encodes the formant amplitudes, frequencies, and their associated bandwidths. Additionally, the pitch and V/UV determinations are made and encoded. Various methods exist for tracking and extracting formant information [50, 83, 85, 112]. More recent efforts [4, 50, 83, 88] have determined that spectral analysis by linear prediction is the most accurate method for formant extraction. Equations A3-37 through A3-40 describing the operation of LPC vocoders result in an error expression for the difference between the actual speech signal, s(n) and the sampled speech signal, s(n). Equation A3-44 is a set of simultaneous, linear autocorrelation equations. This set of equations result in a spectral analysis of the speech signal. From this spectral signal, the formant frequencies and amplitudes are determined by simple-peak picking methods. More sophisticated vocoders employ additional methods [132] for extracting formants when they tend to merge together or when "extra" peaks are identified as possible formants. These methods are usually implemented in software to effect decision-making options. Figure A3-6 and A3-7 show example block diagrams of formant vocoders.

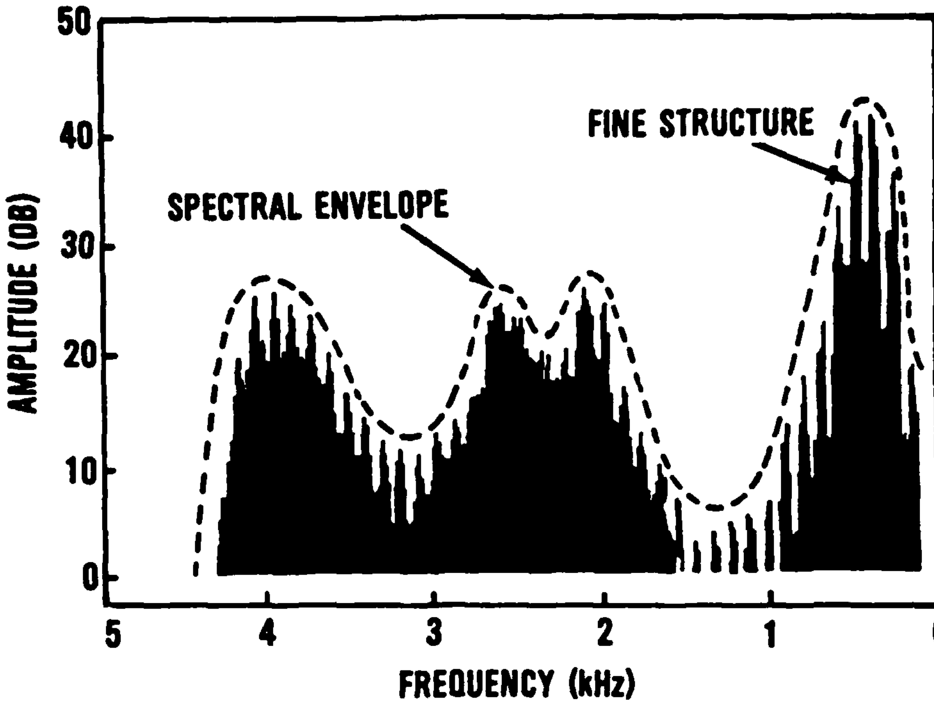



Figure A3-5. Logarithmic spectrum of "EE" sound in procEEdings. Spectral envelope shows four peaks or formants. The fine structure corresponds to a Fundamental frequency of about 110 HZ.
(Taken from ref. 115, p. 722)

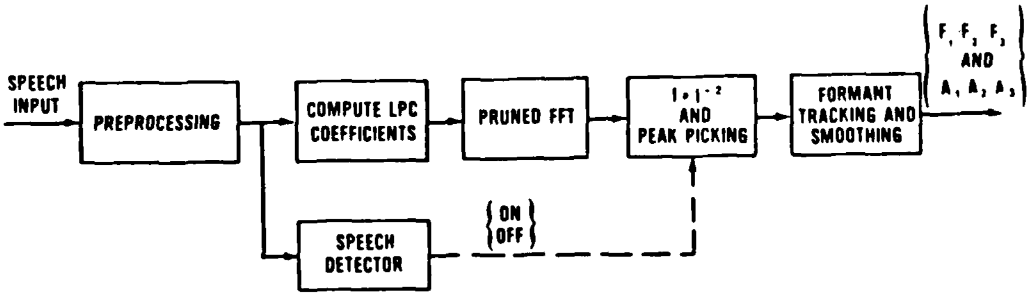



Figure A3-6. Block diagram of the procedure for determining formant frequencies and amplitudes.
(Taken from ref. 132, p. 346)

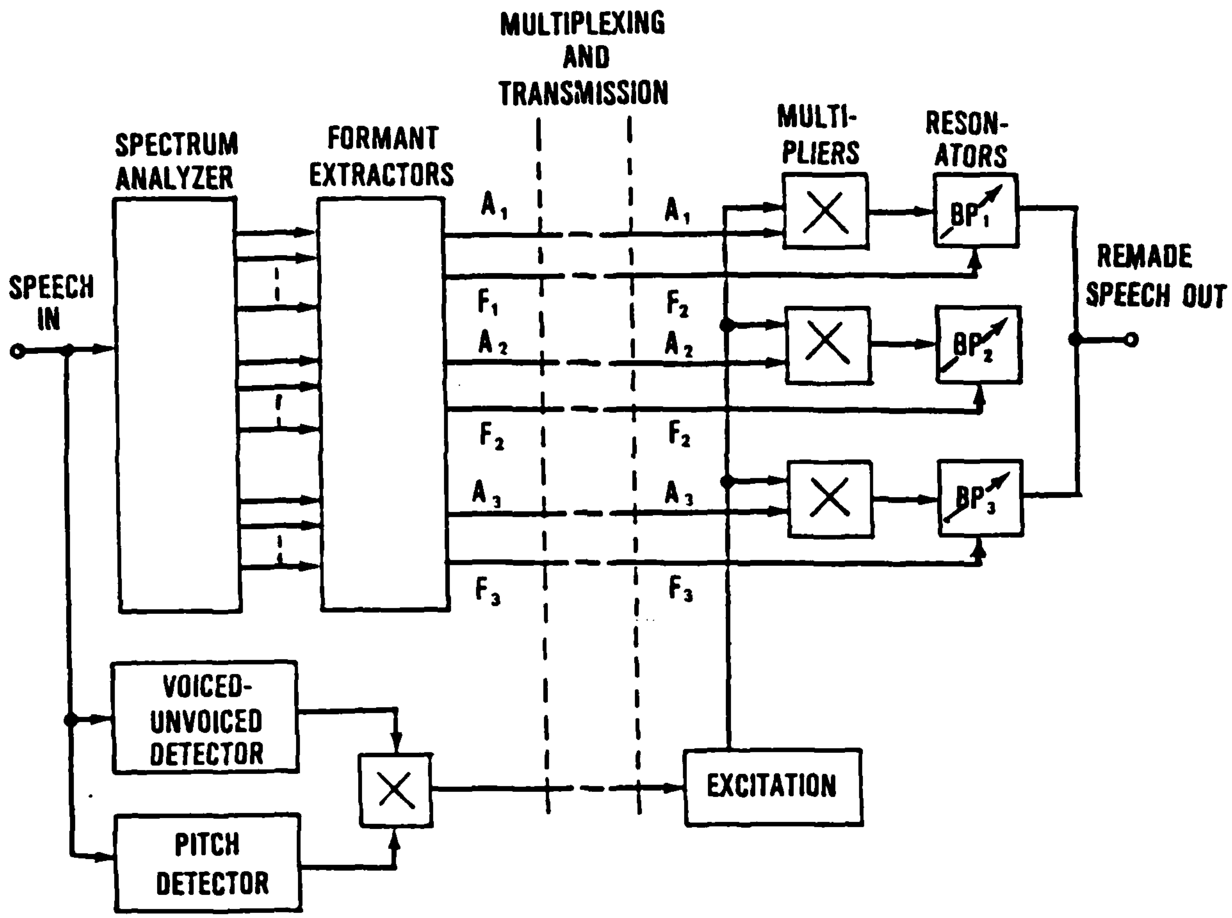



Figure A3-7. Formant vocoder using parallel synthesis.
(Taken from ref. 115, p. 725)

A second method is based on spectral moments, such as the centroid of the signal spectrum in a formant frequency band (116). The mean frequency can be measured by dividing the speech band into sub-bands with filters and measuring the amplitude, $a_n$, in each. The mean frequency is then:

$$f_{mean} = \frac{\sum_{n}^{N} a_n f_n}{\sum_{n}^{N} a_n} . \qquad (A3-7)$$

where N is the number of subbands utilized. This can be performed in the time domain by:

$$f_{mean} = \frac{1}{2\pi} \frac{\overline{\left| \frac{d}{dt} s(t) \right|}}{\left| \overline{s(t)} \right|} , \qquad (A3-8)$$

where the bars indicate the long-time averages.

The problem with formant vocoders of this type is that usually the formant sub-bands overlap.

Analysis-by-synthesis vocoders (7) are a highly specific form of formant vocoder not utilizing the previously described analysis methods. These vocoders iteratively generate artificial spectra which are matched to the windowed segment of the speech spectrum. After matching, the formant characteristics of the spectrum generator are taken as those of the actual speech. For real-time operation, a highly complex, extremely fast computer is necessary in order to generate the spectrum iterates fast enough.

Formant vocoders use the same V/UV and pitch extraction techniques as the channel vocoders. The synthesis system uses the same pulse or noise excitation signals for remodulating the speech signal.

Homomorphic Vocoder

The homomorphic or cepstrum vocoder is an even more complex, sophisticated system of speech analysis-synthesis than those previously mentioned. These systems utilize more recent advances in FFT computation and signal deconvolution to analyze and then synthesize the voice communications. A high-resolution spectral analysis or Fourier transform is computed on a windowed segment (20-40 ms) of speech (42). Within this segment, the speech waveform is very nearly stationary, thereby allowing the application of the Fourier transform resulting in the spectrum amplitude, as in the channel vocoder. The logarithm of the spectrum amplitude is then computed. This process changes the speech waveform from a convolution of two functions into a product of two functions and finally into a sum of two functions. (The Z-transform could be utilized as readily as the Fourier transform.) The speech signal is described by:

$$s(nt) = v(nt)*e(nt), \qquad \text{(A3-9)}$$

where $s(nt)$ is the speech waveform, $v(nt)$ is the vocal tract impulse response, and $e(nt)$ is the excitation function. The spectrum is given by:

$$S(f) = V(f)E(f). \qquad \text{(A3-10)}$$

The spectrum log magnitude, $\hat{S}(f)$, is given by:

$$\hat{S}(f) = \ln[S(f)] \tag{A3-11}$$

Inserting (A3-10) yields:

$$\hat{S}(f) = \ln[V(f)E(f)] \tag{A3-12}$$

which expands into:

$$\hat{S}(f) = \ln[V(f)] + \ln[E(f)] \tag{A3-13}$$

The log magnitude of the spectrum is then transformed back into the time domain (inverse transform). This results in the ceptsrum, C(nt), a signal where the envelope function (vocal-tract response) is concentrated in the low-time values and the excitation (pitch) appears as a periodic set of lines. The cepstrum is given by:

$$c(nt) = F^{-1}[\hat{S}(f)] \ , \tag{A3-14}$$

therefore,

$$c(nt) = F^{-1}(\ln[V(f)]) + F^{-1}(\ln[E(f)]) \ . \tag{A3-15}$$

If a low-time window is multiplied with the cepstrum, a smoothed envelope is the result (cepstral smoothing). This smoothed envelope is easily quantized and transmitted to the synthesizer. Since the pitch appears as a series of pulses, the period is reasonably easy to determine. This process is depicted in Figure A3-8. Figure A3-9 shows a block diagram of a cepstrum vocoder.

Pattern-Matching Vocoder

The pattern-matching vocoder is another spectrum analysis system. It has stored within it a series of spectral patterns which

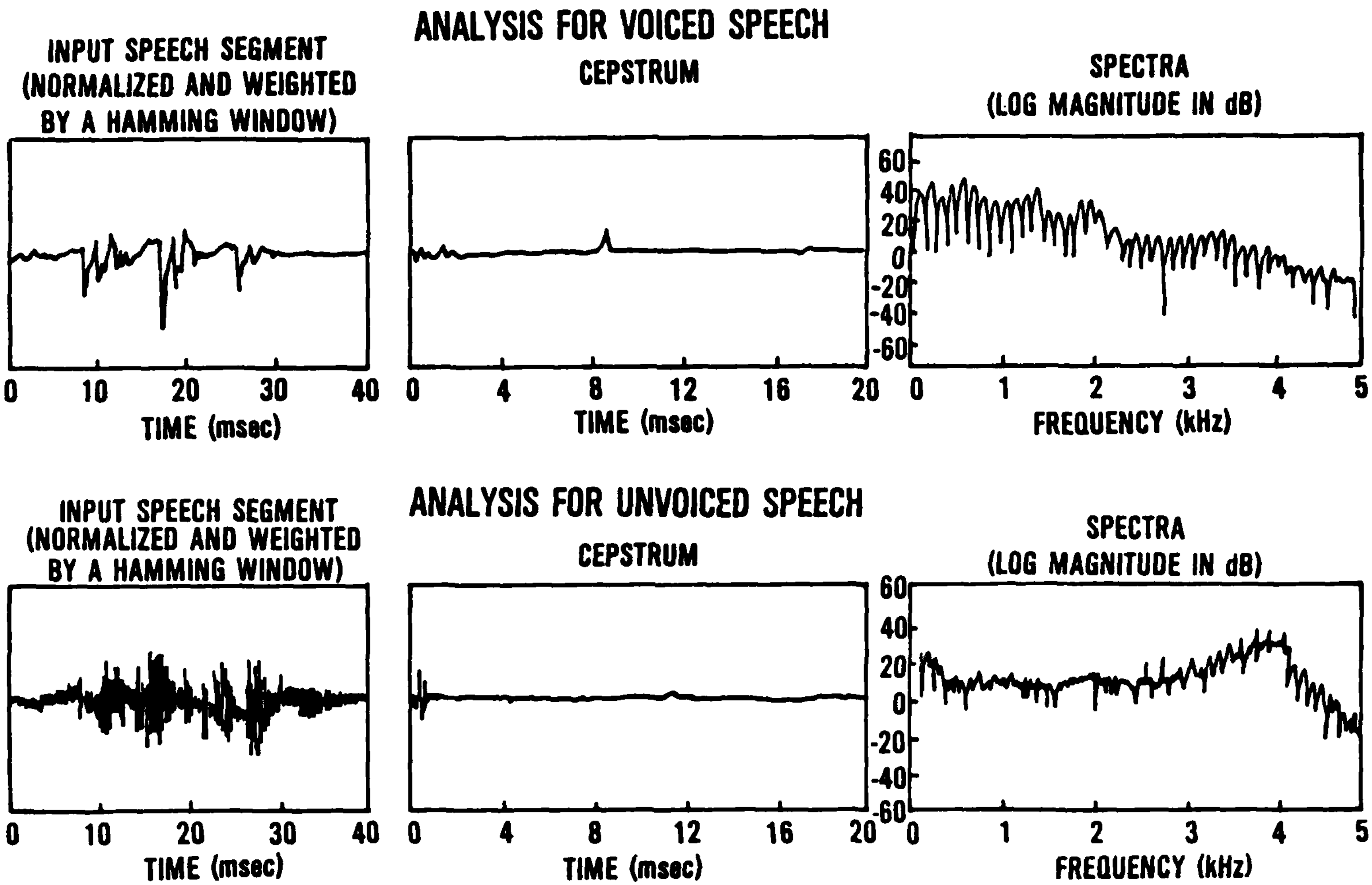



Figure A3-8. Homomorphic analysis for voiced and unvoiced speech.
(Taken from ref. 104, page 690)

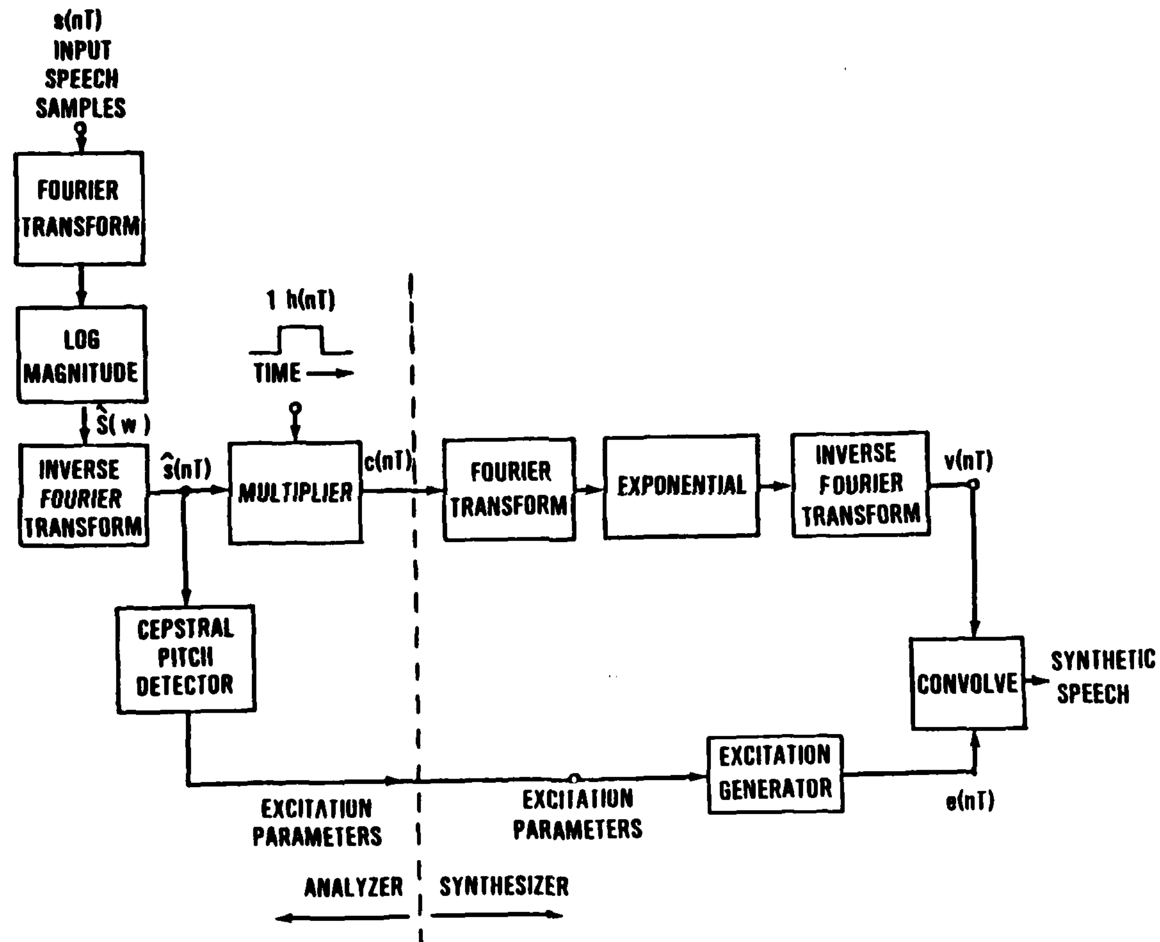



Figure A3-9. Block diagram of a homomorphic vocoder.
(Taken from ref. 42, p. 1643)

match the number of "just-discernible" sounds. The number of these patterns differs from system to system according to how the designer determined what constitutes the just-discernible sounds or according to what he determines is the minimum number of sound patterns necessary to accurately recreate the speech sounds. The pattern matching in today's digital systems is generally performed by point-by-point subtraction. The patterns to be matched can be generated by any number of other schemes. A channel-type vocoder system can be employed to generate smaller frequency elements each of which is a segment of the Fourier transform. Or a formant analysis may be performed with the formant trajectories compared to idealized formants. Cepstral analysis could also be employed. Once the speech is analyzed, it is compared to the patterns in memory. For a pattern to be a match, the error must satisfy some minimum error distance measure. For transmission, the memory location of the "best guess" pattern is transmitted. Additionally, the pitch and V/UV data must be determined and transmitted. At the synthesizer, the address is used to retrieve the "correct" pattern. This pattern is then modulated with the pitch and V/UV source to regenerate the speech sounds. With the fast microprocessors available today, this method is becoming more appealing.

## Phase Vocoder

In the phase vocoder, the speech signal is represented by its complex, short-time Fourier transform or in other words by its short-time amplitude and phase spectra (35, 111). The system uses a bank of adjacent bandpass filters (channels) to perform the Fourier

analysis. After filtering, the channels are recombined with an essentially insubstantial degradation. This is shown in Figure A3-10. The output of the n-th filter is $f_n(t)$. The reconstructed, approximate signal is given by:

$$f(t) \cong \sum_{n=1}^{M} f_n(t). \tag{A3-16}$$

The n-th filter has an impulse response given by:

$$g_n(t) = h(t) \cos w_n t , \tag{A3-17}$$

where h(t) is the impulse response of a physically-realizable low-pass filter. The filter output is:

$$f_n(t) = f(t) * g_n(t), \tag{A3-18}$$

which is expanded into:

$$f_n(t) = \int_{-\infty}^{t} f(\lambda)h(t-\lambda) \cos [w_n(t-\lambda)]d\lambda \tag{A3-19}$$

or

$$f_n(t) = \mathrm{Re}(e^{jw_n t} \int_{-\infty}^{t} f(\lambda)h(t-\lambda)e^{-jw_n\lambda}d\lambda) \tag{A3-20}$$

where the integral in (A3-20) is the short-time Fourier transform. The transform can be expressed in terms of its amplitude and phase as

$$f_n(t) = \left| F(w_n,t) \right| \cos [w_n t + \psi(w_n,t)] , \tag{A3-21}$$

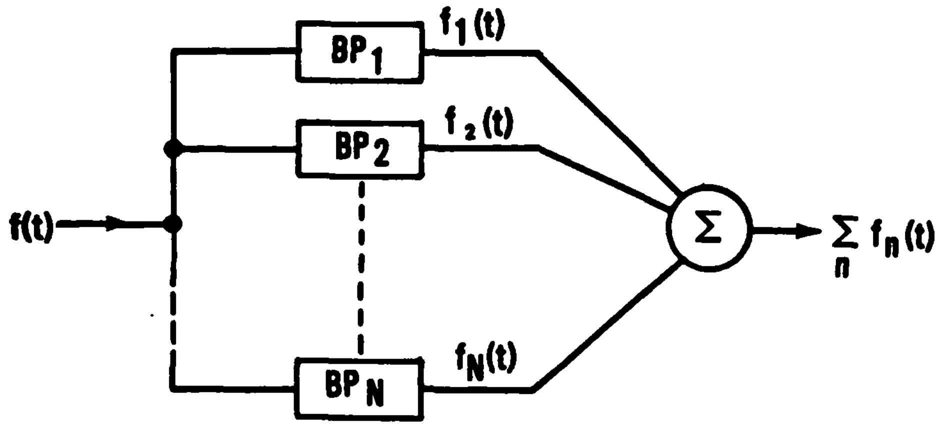



Figure A3-10. Filtering of speech by adjacent band-pass filters.
(Taken from ref. 35, p. 1494)

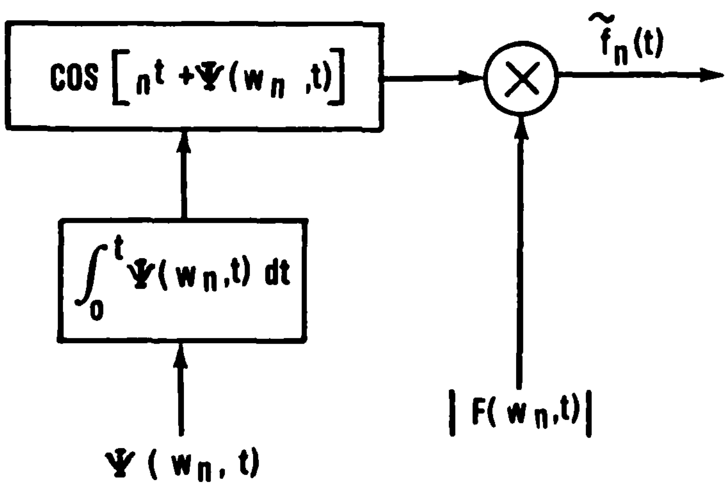



Figure A3-11. Speech synthesis based on the short-time amplitude and phase derivative spectra.
(Taken from ref. 35, p. 1497)

where $F(w_n,t)$ is the complex short-time transform and $\psi(w_n,t)$ is the short-time phase spectrum. $|F(w_n,t)|$ can be bandlimited to 20 or 30 Hz without perceptual distortion. $\psi(w_n,t)$ is unbounded and therefore not suitable for transmission. The time derivative, $\dot{\psi}(w_n,t)$, is computed for transmission. Now the signal can be approximated as:

$$\tilde{f}_n(t) = |F(w_n,t)| \cos [w_n t + \tilde{\psi}(w_n,t)], \tag{A3-22}$$

where

$$\tilde{\psi}(w_n,t) = \int_{-\infty}^{t} \dot{\psi}(w_n,t)\, dt. \tag{A3-23}$$

(A3-23) shows how $(w_n,t)$ is recovered to within the value of an additive constant. Because the ear is relatively insensitive to phase this phase error constant poses no serious problem.

The synthesizer reconstructs the signal by summing the outputs of n oscillators modulated in phase and amplitude from bandlimited versions of $|F(w_n,t)|$ and $\dot{\psi}(w_n,t)$. This process is shown in Figure A3-11.

A computer implementation is shown in Figure A3-12. The mathematical process utilizes the real and imaginary components of the complex spectrum

$$F(w_n,t) = a(w_n,t) - jb(w_n,t) , \tag{A3-24}$$

where

$$a(w_n,t) = \int_{-\infty}^{t} f(\lambda)h(t-\lambda) \cos (w_n\lambda)d\lambda \tag{A3-25}$$

and

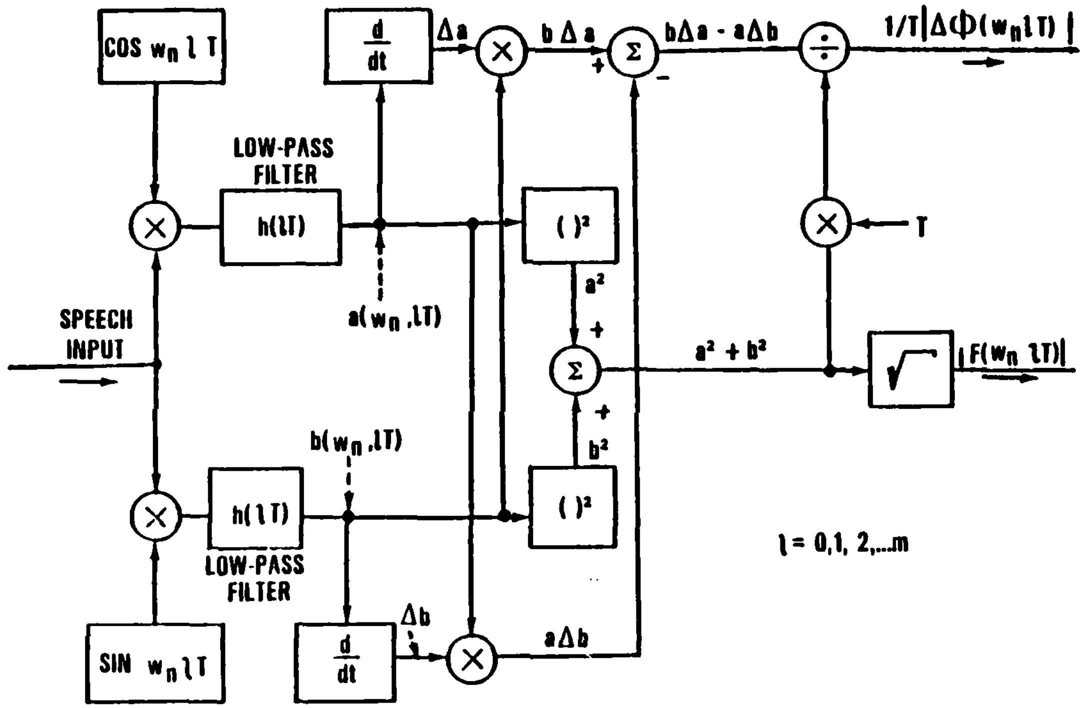



Figure A3-12. Programmed operations for extracting $|F(w_n,t)|$ and $\dot{\psi}(w_n,t)$. (Taken from ref. 35, p. 1498)

$$b(w_n,t) = \int_{-\infty}^{t} f(\lambda)h(t-\lambda)\sin(w_n\lambda)d\lambda \qquad \text{(A3-26)}$$

Now the amplitude and phase derivatives are given by:

$$|F(w_n,t)| = (a^2 + b^2)^{1/2} \qquad \text{(A3-27)}$$

and

$$\dot{\psi}(w_n,t) = \left(\frac{a\dot{b}-b\dot{a}}{a^2+b^2}\right) \qquad \text{(A3-28)}$$

For computer implementation these equations are given by:

$$a(w_n,t) = T\sum_{l=0}^{m} f(lt)[\cos(w_n lt)]h(mT-lT) \qquad \text{(A3-29)}$$

and

$$b(w_n,t) = T\sum_{l=0}^{m} f(lt)[\sin(w_n lt)]h(mT-lT)\ , \qquad \text{(A3-30)}$$

where T is the sampling interval. The derivatives are computed as:

$$\Delta a = a[w_n,(m+1)T] - a[w_n,mT] \qquad \text{(A3-31)}$$

and

$$\Delta b = b[w_n,(m+1)T] - b[w_n,mT]\ . \qquad \text{(A3-32)}$$

The magnitude and phase derivatives in discrete form are:

$$|F(w_n,mT)| = (a^2+b^2)^{1/2} \qquad \text{(A3-33)}$$

and

$$\frac{\Delta\psi}{T}[w_n, mT] = \frac{1}{T}\left(\frac{b\Delta a - a\Delta b}{a^2+b^2}\right) \quad \text{(A3-34)}$$

The computer resynthesis for one channel is:

$$\tilde{f}(mt) = |F(w_n, mT)| \cos\left(w_n mT + T \sum_{l=0}^{m} \frac{\Delta\psi(w_m, lT)}{T}\right) . \quad \text{(A3-35)}$$

This is summed for all channels to recover the speech signal. This method eliminates the need for pitch and V/UV decision extraction.

<u>Linear Predictive Coding Vocoders</u>

Linear Predictive Coding (LPC) is, at this time, the most common method of vocoder implementation. The LPC process is usually based on autocorrelation analysis. The speech signal is highly repetitive and redundant in its features. As indicated previously, the vocal tract is a slowly varying system; this is what gives rise to the repetitiveness of speech. This characteristic is easily exploited, in that adjacent segments of the signal are highly correlated. This means that given n past samples of the waveform the next segment can be predicted with generally a high degree of accuracy. Increasing <u>n</u> tends to improve the prediction accuracy. These predictors are weighted values recomputed every 20-40 ms. This type of analysis is generally attempting to model the vocal tract. The most general form of this model is:

$$H(z) = G \frac{1+\sum_{l=1}^{q} b_l z^{-l}}{1-\sum_{k=1}^{p} a_k z^{-k}} . \qquad (A3\text{-}36)$$

It has been found (95) that it is sufficiently accurate to allow the numerator of H(z) to equal G, a constant gain. This results in an all-pole model.

In general terms an LPC vocoder generates the predictor coefficients. The coefficients are then used in a correlation procedure and matched with the incoming segment of speech just predicted. This results in the generation of an error signal. The next set of predictors are generated in order to minimize the error.

In LPC systems, speech is modeled by an all pole filter, H(z). Thw filter transfer function is given by:

$$H(z) = \frac{G}{1-\sum_{k=1}^{p} a_k Z^{-k}} . \qquad (A3\text{-}37)$$

The frequency domain and time domain models for this process are shown in Figure A3-13. This model assumes that a frame or window of speech can be expressed as:

$$s(n) = \sum_{k=1}^{p} a_k s(n-k)+U_n \quad , \qquad (A3\text{-}38)$$

where p is the number of poles, $U_n$ is the appropriate input excitation, and the $a_k$'s are the predictor coefficients characterizing the filter. To generate the speech signal knowledge of the pitch, filter

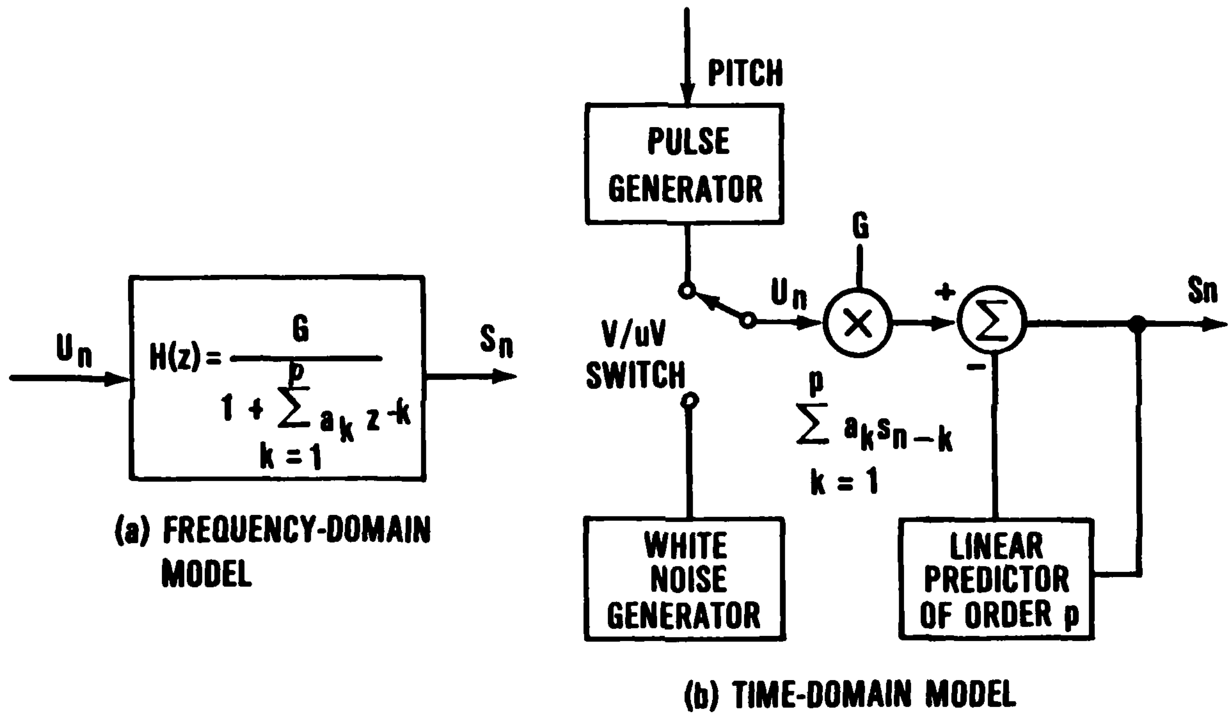



Figure A3-13. Discrete model of speech production as used in linear prediction. (Taken from ref. 110, p. 1694)

parameters, and the gain of the filter (amplitude of input excitation) in each frame is needed (11). Filters with various numbers of poles are used, with eight to sixteen being the most common.

In (A3-38), $U_n$ is zero except for one sample at the beginning of each pitch period. Thus, the equation becomes:

$$\hat{s}(n) = \sum_{k=1}^{p} a_k s(n-k) \quad . \tag{A3-39}$$

Now if the model were perfect, the speech samples, S(n), would be completely predictable. This perfect model does not exist, therefore, it is necessary to define an error, E(n), between s(n), the sampled speech, and s(n), the predicted speech. This error is given as:

$$E(n) = s(n) - \hat{s}(n) = s(n) - \sum_{k=1}^{p} a_k s(n-k) \quad . \tag{A3-40}$$

The mean square error is given by:

$$E_T = \langle E(n)^2 \rangle = \sum_{n=1}^{\infty} [s(n) - \sum_{k=1}^{p} a_k s(n-k)]^2 \quad . \tag{A3-41}$$

The $a_k$'s are chosen so as to minimize this error. This can be done by computing:

$$\frac{\partial \langle E(n)^2 \rangle}{\partial a_j} = 0 \quad , j = 1, 2, \ldots , p \; , \tag{A3-42}$$

yielding the set of equations:

$$\sum_{k=1}^{p} a_k \sum_{n=1}^{\infty} s(n-k)s(n-j) = \sum_{n=1}^{\infty} s(n)s(n-j), \quad j = 1, 2, \ldots, p \,. \tag{A3-43}$$

The right side of this equation constitutes an autocorrelation function, $R_j$, and (A3-43) can be expressed as:

$$\sum_{k=1}^{p} a_k R(j-k) = R_j \,, \quad 1 \leq j \leq p. \tag{A3-44}$$

This set of equations can be solved recursively as follows:

$$E_0 = R_0 \tag{A3-45a}$$

$$K_j = \frac{-(R_j + \sum_{n=1}^{j-1} a_n^{(j-1)} R_{j-k})}{E_{j-1}} \,, \tag{A3-45b}$$

$$a_j^{(j)} = k_j,$$

$$a_n^{(j)} = a_n^{(j-1)} + k_j a_{j-k}^{(j-1)}, \quad 1 \leq n \leq j-1, \tag{A3-45c}$$

$$E_j = (1-k_j^2)E_{j-1} \,, \tag{A3-45d}$$

The final solution is given by:

$$a_n = a_n^{(p)} \,, \; 1 \leq n \leq p \,. \tag{A3-45e}$$

The $k_j$'s are reflection coefficients (or partial correlation coefficients). By expanding the squared terms in (A3-41) and using (A3-44), the minimum error is given by:

$$E_p = R_0 + \sum_{k=1}^{p} a_k R_k \ . \tag{A3-46}$$

Given the above derivation for error and predictor coefficients, alternate sets of transmission parameters can be derived. Given below is a list of the possible sets of parameters for characterizing uniquely the linear prediction filter H(z) (77). The denominator of H(z) is an inverse filter A(z) given by:

$$A(z) = 1 + \sum_{k=1}^{p} a_k z^{-k} \ . \tag{A3-47}$$

The various parameters suitable for transmission are the:

1. Impulse response of A(z), i.e., the predictor coefficients $a_k$, $1 \leq k \leq p$,

2. Impulse response of the all-pole model, $h_k$, $0 \leq k \leq p$, easily obtained by long division,

3. Autocorrelation coefficients of $(a_k/G)$ given by:

$$b_n = \frac{1}{G^2} \sum_{n=0}^{p-|i|} a_n a_{n+j} \ , \ a_0 = 1, \ 0 \leq j \leq p, \tag{A3-48}$$

4. Autocorrelation coefficients of $(h_k)$ (partial correlation coefficients) given by:

$$r_j = \sum_{n=0}^{\infty} h_n h_{n+j} \ , \ 0 \leq j \leq p \ , \tag{A3-49}$$

where

$$r_j = R_j \text{ in (A3-44) for } 0 \leq j \leq p \ ,$$

5. Spectral coefficients of A(z)/G, $P_j$, $0 \leq j \leq p$ (or equivalently, spectral coefficients of H(z), $1/P_j$ given by:

$$P_j = b_0 + 2 \sum_{n=1}^{p} b_n \cos \frac{2\pi jn}{2p+1} \ , \ 0 \leq j \leq p \tag{A3-50}$$

(this results in a Linear Prediction Channel Vocoder),

6. Cepstral coefficients (log-area-ratio coefficients) of A(z), $c_k$, $1 \leq k \leq p$ (or equivalently cepstral coefficients of H(z)/G, -Ck given by:

$$C_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log A(e^{jw}) e^{jnw} dw, \tag{A3-51}$$

or in digital form

$$C_k = a_k - \sum_{m=1}^{k-1} \frac{m}{k} C_m a_{k-m}, \quad 1 \leq k \leq p \ , \tag{A3-52}$$

7. Poles of H(z) (or zeros of A(z)),

8. Reflection coefficients, $K_j$, $1 \leq j \leq p$ or simple transformations thereof, i.e., area coefficients given by:

$$A_k = A_{k+1} \frac{1+k_j}{1-k_j} \ , \ A_{p+1}=1 \ , \ 1 \leq j \leq p \tag{A3-53}$$

(The reflection coefficients are an intermediate product of the error minimization but may be computed directly by:

$$k_j = a_j^{(j)} \tag{A3-54}$$

$$a_n^{(j-1)} = \frac{a_n^{(j)} - a_n^{(j)}\, a_{j-n}^{(j)}}{1-k_j^2}\ ,\ 1 \leq n \leq j-1 \qquad \text{(A3-55)}$$

where j takes the values p, p-1,...,1. Initially

$$a_n^{(p)} = a_n,\ 1 \leq j \leq p\ .)\ , \qquad \text{(A3-56)}$$

and

9. The error coefficients. (If the analysis and synthesis systems are guaranteed to start at the same state, a measurement of the signal matching error can be used to adjust the synthesizer.)

Figure A3-14 shows a block diagram of a linear predictive coding system. Figure A3-15 shows a pole-zero predicter block diagram. The zero portion is not implemented. Pitch and V/UV decisions must also be supplied by the analyzer. Data transmission rates currently vary from about 1.2 kbps to 10 kbps depending upon implementation.

## Spectral Evelope Estimation

The Spectral Envelope Estimation (SEE) vocoder is similar to the homomorphic vocoder in that it utilizes the log magnitude of the Fourier transform. This is shown in Figure A3-16. It starts with the assumption that speech can be modeled by:

$$S(f) = E(f,T)V(f)\ , \qquad \text{(A3-57)}$$

where E(F,T) is a unity amplitude impulse train of period 1/T (the pitch period) and $|S(f)|$ is a sampling of the vocal tract response $|V(f)|$ at the points f=k/T for k=1,2,... (99). Interpolating

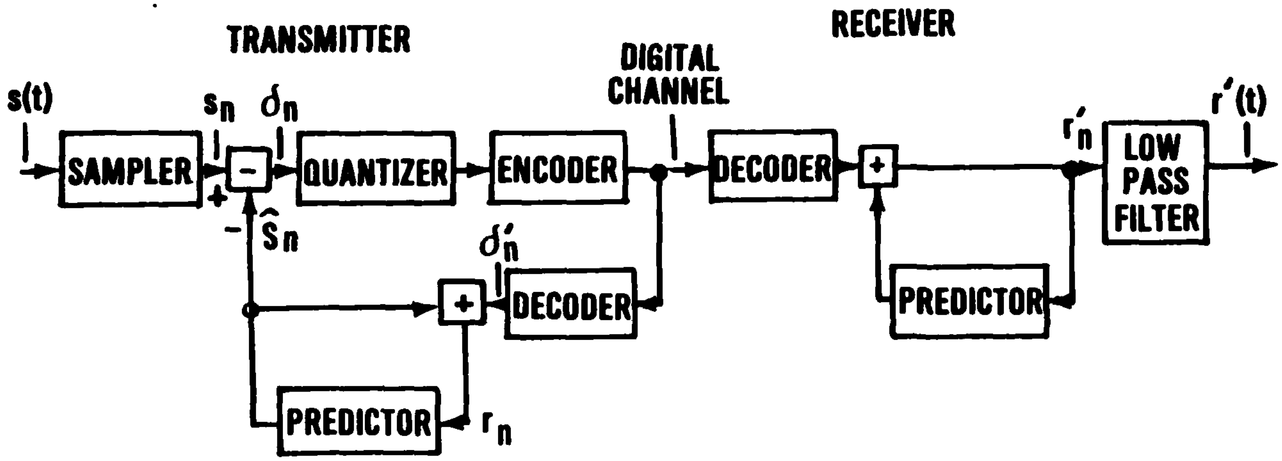



Figure A3-14. Block diagram of a predictive coding system.
(Taken from ref. 3, p. 231)

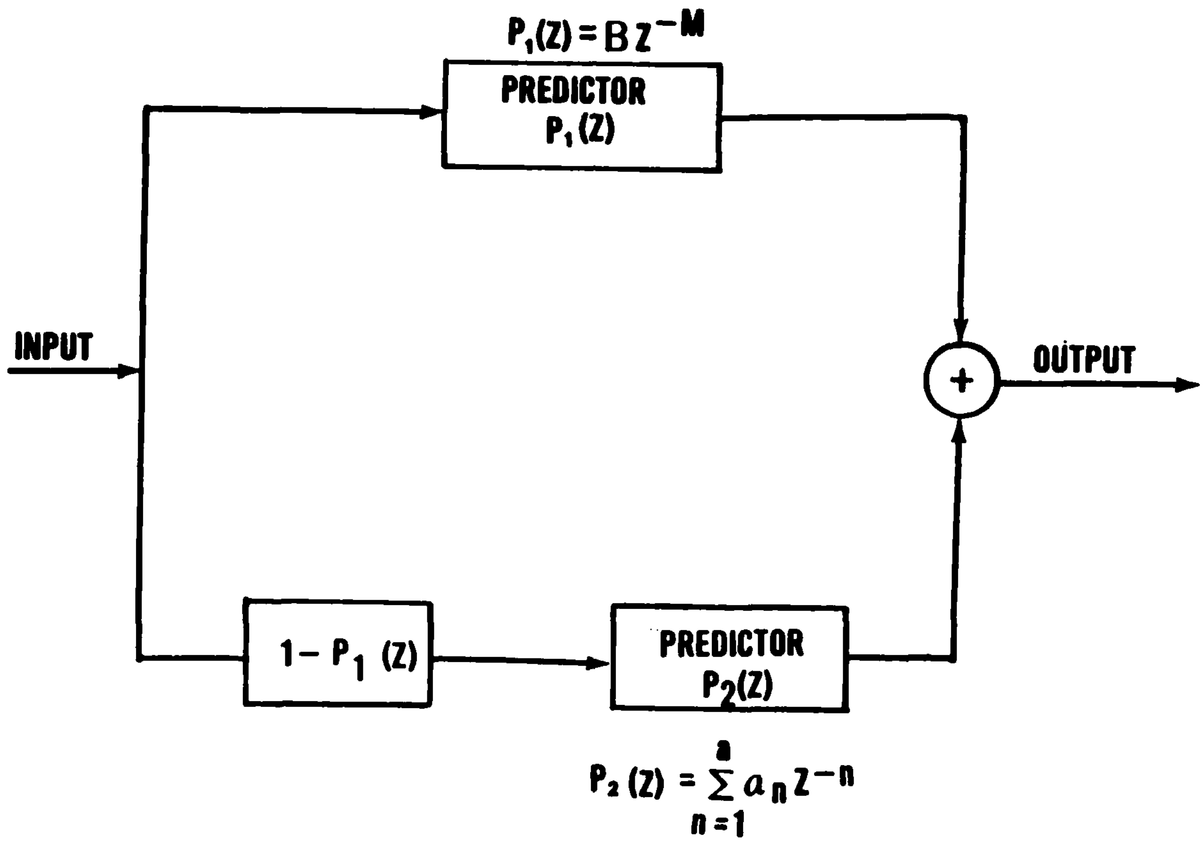



Figure A3-15. Block diagram of the predictor for speech signals.
(Taken from ref. 3, p. 234)

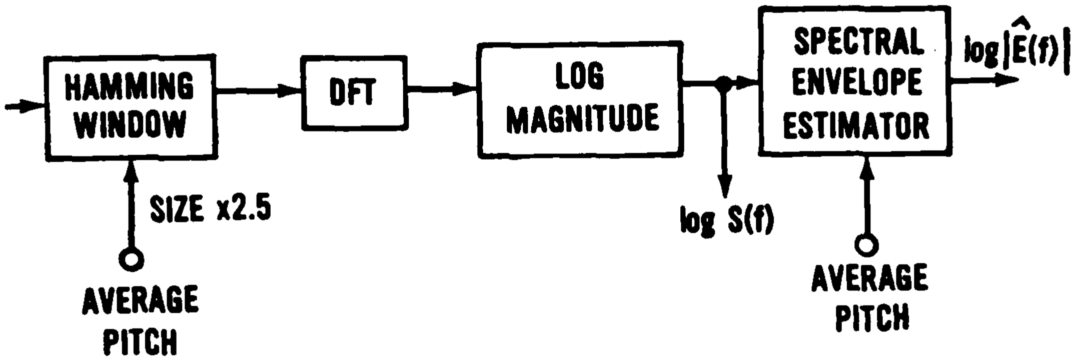



Figure A3-16. Spectral analyzer structure.
(Taken from ref. 99, p. 787)

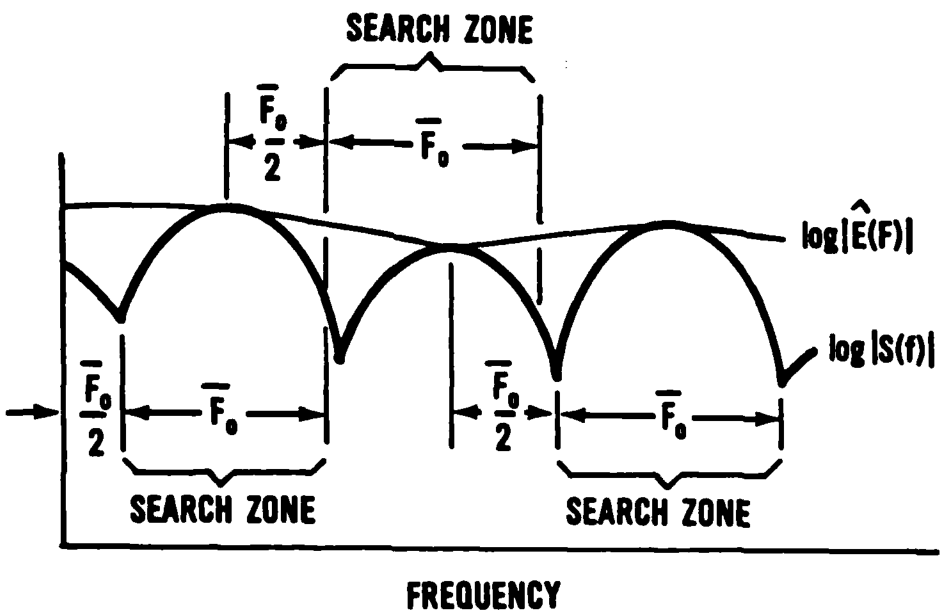



Figure A3-17. Spectral envelope estimator.
(Taken from ref. 99, p. 787)

between the sample points $|S(k/T)|$ will provide the spectral envelope estimation, $|\bar{S}(f)|$. If enough samples are taken, i.e., if $|V(f)|$ is reasonably smooth relative to 1/T, the envelope estimate will approximate the ideal vocal tract response,

$$|V(f)| \cong |\bar{S}(f)| . \qquad \text{(A3-58)}$$

In order to use $|\bar{S}(f)|$ as the estimate, the location of the samples, $|S(k/T)|$, must be reasonably accurate. To locate the peaks of $|S(f)|$, the speech is assumed stationary within a frame or windowed segment of the speech wave. The procedure for finding these peaks is shown in Figure A3-17. The system samples the pitch detection signal and maintains a short-term average of the pitch, $\overline{F_0}$, a characteristics of each individual speaker. This average is given by:

$$\overline{F_0} = \frac{1}{\text{Avg. Pitch Period}} , \qquad \text{(A3-59a)}$$

with

$$k=1 \text{ and } F_0 = 0 . \qquad \text{(A3-59b)}$$

Next the windowed segment of $|S(f)|$ is searched for $f_k$ such that $|S(f)|$ is maximized where:

$$f_{k-1} + \frac{1}{2} (\overline{F_0}) \leq f \leq f_{k+1} + \frac{3}{2} (\overline{F_0}). \qquad \text{(A3-59c)}$$

This is repeated until the entire segment is covered. If there is a significant difference between the talker's pitch and the average pitch, confusion could occur between the peaks of the spectral lines and the spurious sidelobes caused by windowing the speech.

Once the sampling is completed, the estimate of the entire waveform of $|\bar{S}(f)|$ is needed. This is accomplished by interpolating between the samples $|S(f_k)|$. Since a talker's pitch varies slowly only over approximately an octave, the average pitch is assumed to be reasonably accurate so that all samples are at or near the desired peaks and linear interpolation between the samples can be used.

Now the spectral envelope estimation is the estimate of the vocal tract (filter) response given as:

$$|\bar{S}(f)| = |\hat{H}(f)|, \tag{A3-60}$$

and $|\bar{S}(f)|$ can be found from

$$\log|\bar{S}(f)| = \begin{cases} \log|S(f_1)| & 0<f<f_1 \\ \log|S(f_k)| & f=f_k,\ k>1 \\ \dfrac{f_k-f}{f_k-f_{k-1}}\log|S(f_{k-1})| + \dfrac{f-f_{k-1}}{f_k-f_{k-1}}\log|S(f_k)|, & f_{k-1}<f<f_k,\ k>1 \end{cases} \tag{A3-61}$$

The above derivation of the spectral envelope estimator is based on the assumption that the harmonics of the periodic impulse source sample the frequency response of the vocal tract filter. The spectral envelope estimator can also be viewed as an adaptive channel (filter bank) vocoder analyzer or an improvement on the homomorphic spectral analyzer. Each iteration of the spectral

line-finding heuristic selects the point in the short-term (windowed DFT) spectrum centered on the next spectral line. Since each point of a short-term spectrum is equivalent to the output of a filter whose characteristics are determined by the time window, the procedure selects an analysis filter bank that has exactly one filter per spectral line with each filter centered on the corresponding spectral line. The linear interpolation between spectral points essentially only creates a smooth (spectral) waveform so that the coder need not know about or transmit the locations of the spectral lines (99).

The system must also provide the actual pitch and the V/UV decision for transmission along with the SEE. Figure A3-18 shows the block diagram of an SEE Vocoder.

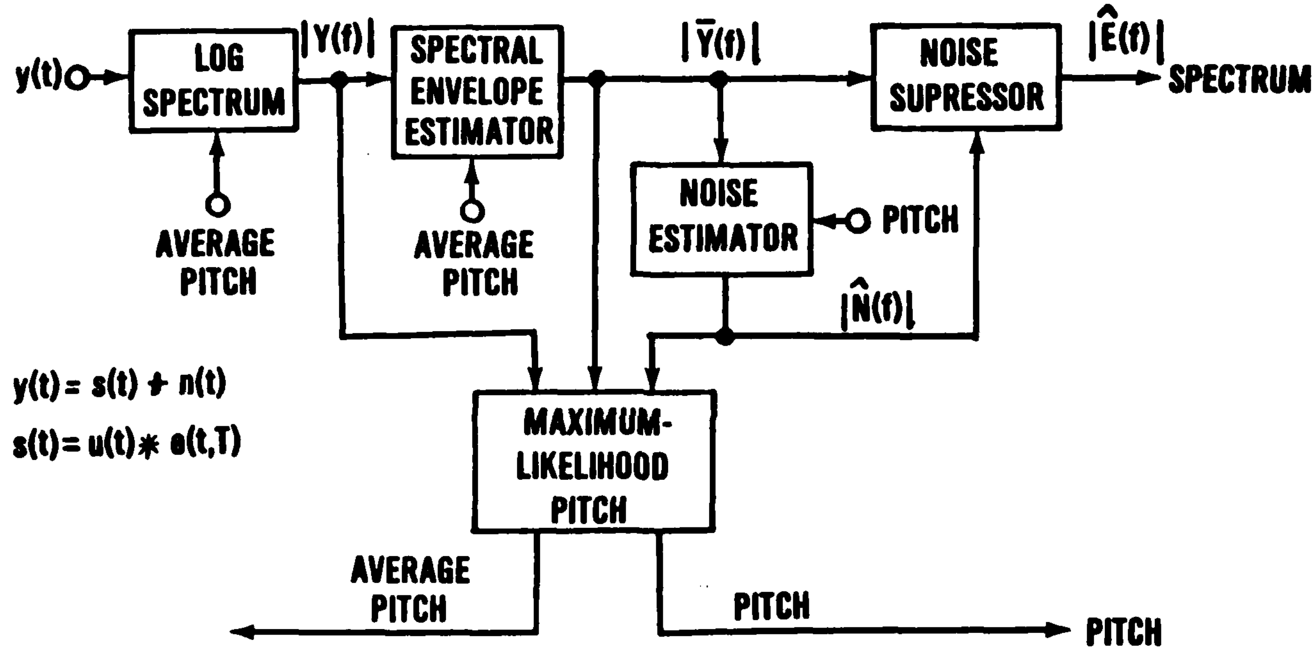



Figure A3-18. Block diagram of spectral envelope estimation vocoder.
(Taken from ref. 99, 791)

END

FILMED

4-85

DTIC